Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 37, NUMBER 1 JANUARY, 1947

Recent Developments in Lenses for Aerial Photography

R. KINGSLAKE
Eastman Kodak Company
(Received October 16, 1946)

IN a previous publication[1] the development of lenses for use in aerial photography was traced from the first beginnings of the subject in 1856 up to the commencement of World War II. It is the purpose of the present paper to continue this story through the war period, and to indicate the lines along which most progress has occurred, at least insofar as the United States is concerned. A very notable parallel development of aerial lenses has taken place in England, and it appears also that some novel and extreme new types of high aperture and wide angle lenses were developed in Germany both immediately before and during the war, but these will not be considered here.

Fortunately the United States was not caught without optical glass as it was in 1917. The principal manufacturers, namely, Bausch and Lomb, Pittsburgh Plate Glass Company, Corning Glass Company, Hayward Optical Glass Company, and the National Bureau of Standards, performed a magnificent service to the country by multiplying their output of optical glass many times and by initiating the manufacture of new types of glass which were previously available only from Europe. The new Kodak high index glass was also used in most of the aerial lenses manufactured by the Eastman Kodak Company. As a result of this fortunate situation, although there were some periods of shortage, especially in the production of disks suitable for very large lenses, the needs of the optical industry were met, and by the end of the war the manufacturers' capacity was actually greater than the need. It is not generally realized, and it cannot be too strongly stressed, how the tremendous efforts of the American optical glass manufacturer made possible the immense number of aerial lenses, not to mention the hundreds of different kinds of telescopes and fire-control instruments, that were needed during the war.

From the very beginning of the war, Canada balsam was replaced by a methacrylate plastic cement to resist high and low temperatures, and the glass-air surfaces of all aerial lenses were coated with an antireflection film. This was of soft calcium fluoride at first, but later hard coatings were developed that offer as much resistance to abrasion as the glass itself.

Aerial camera lenses fall naturally into three groups, namely mapping and survey lenses, night-photo lenses, and lenses for reconnaissance cameras. As in the previous paper, these three types of lens will be discussed separately.

(A) MAPPING AND SURVEY LENSES

The standard lens for all aerial survey work throughout the war was the well-known Metrogon lens of Bausch and Lomb (based on Richter's "Topogon" lens, DRP 636,167). The production rate of this lens rose to the astonishing peak of

[1] R. Kingslake, J. Opt. Soc. Am. **32**, 129 (1942).

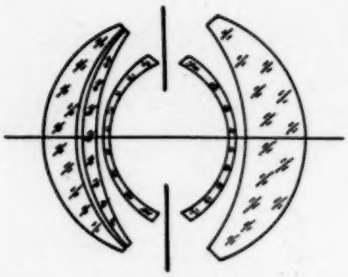

FIG. 1. Rayton's modification of the Metrogon lens. U.S. 2,325,275.

about 1000 lenses a month, in spite of the almost insuperable difficulties involved in its manufacture. Most of these lenses were of 6-inch focal length and f:6.3 aperture, covering a 9″×9″ picture on the Fairchild K-17 camera. A small number of Metrogon lenses were also made in 3-inch, 4-inch, and 12-inch sizes. A common application was in the "Tri-Metrogon" arrangement of three cameras, one pointing vertically downwards and the other two directed obliquely outwards on either side, the lens axes being at 60° to the vertical. In this way the three cameras would make a record of a continuous strip of territory from the left horizon to the right horizon, thus orienting photographs taken over unknown or uncontrolled territory. By taking a succession of these triple photographs as the plane moves forward, the width of the field of a single Metrogon could be about doubled by making use of the two side oblique views in addition to the normal central picture. This equipment was widely used during the war.

There is unfortunately a small residual of zonal distortion in the Metrogon lens, which is large enough to be significant in accurate contour surveying. In an attempt to improve this condition, W. B. Rayton patented a modified design (Fig. 1) in which the front crown lens is split into two closely adjacent positive elements (U.S. 2,325,275). As with so many other new developments, the war came to an end before the necessary tests could be made and the new lens put into production.

(B) NIGHT-PHOTO LENSES

At the start of the war, the only lens suitable for night-photo work with a flash bomb was the Kodak 13½-inch f:3.5 lens originally fitted to the old Fairchild K-12 camera using 8″×10″ cut films, but later adapted to fit the K-19 camera. The design of this lens was improved, and a number were manufactured to the new formula shown in Fig. 2. However, it was soon realized that f:3.5 is inadequate for satisfactory night-photo work, and at the request of the Army Air Corps, the Eastman Kodak Company designed a 7-element f:2.5 lens (Fig. 3), which was manufactured by hundreds in the 12-inch focal length for the K-19 camera (9″×9″), and by thousands in the 7-inch focal length for the K-24 camera (5″×5″). A small number of other experimental 7-inch f:2.5 lenses were made for the N.D.R.C. by the Harvard Observatory Optical Laboratory, some employing plastic elements (made in collaboration with the Polaroid Corporation) and some using only cheap ophthalmic glass with thin elements and weak curves, the purpose being to develop rapid production in a short time.

A large number of night photographs were taken by the services, mainly by means of explosive flash bombs, the camera shutter being operated automatically by a photoelectric cell. Some pictures were taken by the light from an Edgerton-type electrical flash source situated in the airplane itself and synchronized directly with the camera shutter, but the electrical equipment required to give a really bright flash is too bulky and heavy for convenient use, and the ground illumination produced by this flash is still further reduced by the necessity for having the source of light at the airplane instead of relatively close to the ground as with the ordinary flash bomb. However, the extreme brevity of the electrical flash makes sharp photographs possible even

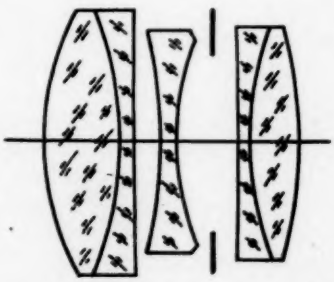

FIG. 2. Kodak 4″ and 13½″ f:3.5 Ektar lens. U.S. 2,279,384 (Altman).

when the airplane is moving in an irregular manner and traveling at a high speed close to the ground.

In an attempt to improve the photographic quality of night photographs, an 8-inch f:1.5 lens was designed by Kodak for use on the 5″×5″ K-24 camera (Fig. 4). It was, however, found that the increased sharpness of definition and the wider field of the 7-inch f:2.5 lens more than offset the 2½ times increase in image brightness of the f:1.5 lens, and as a result only a very few of the f:1.5 lenses were actually manufactured.

(C) RECONNAISSANCE LENSES

By far the greatest use of aerial photography in warfare is in routine reconnaissance flights over the enemy's territory, and the major needs in this field are (a) the best possible definition and (b) means for avoiding the enemy's gunfire and pursuit planes. For many years prior to the war, Colonel G. W. Goddard had advocated the use of long-focus lenses carried by large planes at a great height, but in this as in so many other ways both Britain and America were inadequately prepared by 1940, and for many months thereafter the British found it necessary to equip fast fighter planes with small automatic cameras such as the Williamson F-24, employing an 8-inch Dallmeyer f:2.9 Pentac lens to cover a 5×5-inch field, and perform difficult feats of picture-snatching at low altitude and high speed. By the end of 1942, a copy of this camera, known as the K-24, was being manufactured by the thousand in the United States for the same purpose. Three pre-focused and readily interchangeable lenses were supplied in cones for this camera, namely, (a) Kodak 7-inch f:2.5 Aero Ektar (Fig. 3), (b) Kodak 12-inch f:5.0 Aero-stigmat (Fig. 5), and (c) Bausch and Lomb 20-inch f:5.6 Telephoto (Fig. 6).

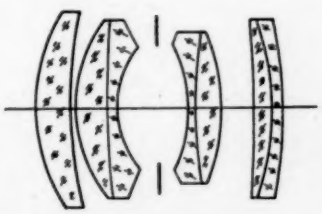

FIG. 3. Kodak 7″ and 12″ f:2.5 Ektar lens. U.S. 2,343,627 (Aklin).

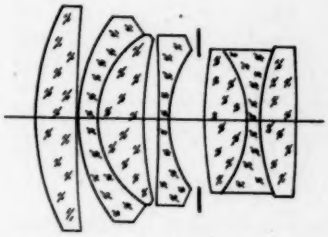

FIG. 4. Kodak 8″ f:1.5 Ektar lens. U.S. 2,397,565 (Schade).

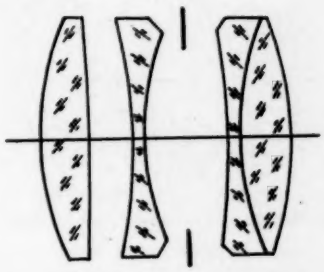

FIG. 5. The Tessar type of construction. Used for 6¾″ f:4.8, 10″ f:4.5, 12″ f:5.0, 24″ f:6.0.

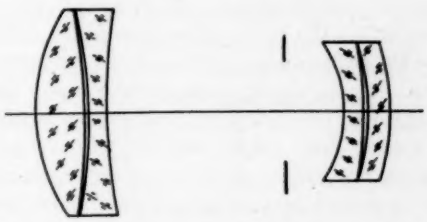

FIG. 6. Bausch and Lomb 20″ f:5.6 Telephoto. U.S. 2,390,387 (Rayton and Hudson).

PHOTOGRAPHIC AIRPLANES

The most popular photographic airplane of the U. S. Army Air Corps was the P-38 "Lightning." The guns were removed from the nose of the plane leaving a compartment large enough to hold the cameras. The pilot was also the photographer, and every possible use was made of the great speed and altitude capabilities of this plane. However, later in the war, as the army began to recognize officially the supreme importance of aerial photographs, and as the extent of military operations increased, it became necessary to equip long range and high flying 2-motored or 4-motored planes with several cameras and to supply a photographer to operate them. These planes were generally un-

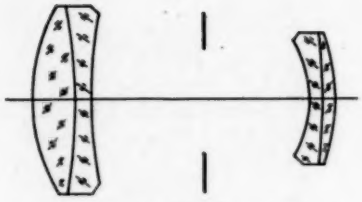

FIG. 7. Wollensak 15" f:5.6 Telephoto.

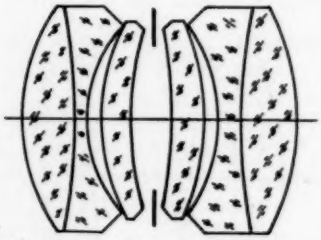

FIG. 8. Bausch and Lomb 8½" f:4 Altimar lens.

armed, relying entirely on speed and altitude for their safety.

The Navy used Multi-engine "Liberator" (PB4Y-1) patrol planes, "Hellcat" (F6F) and "Wildcat" (F4F) fighters, and "Avenger" (TBP) torpedo-bombers especially modified for photography. In the early part of the war Navy photographic reconnaissance of enemy bases in the Pacific was made entirely with "Photo-Liberators." When aircraft carrier operations were increased and conducted within striking range of enemy territory, the "Hellcats" equipped for Tri-Metrogon and the "Wildcats" equipped for vertical and oblique photography were used. The "Avenger" was employed largely for anti-submarine photography.

AERIAL CAMERAS

During the war, the U. S. Army and Navy Air Forces made extensive use of the cameras listed in Table I.

The Navy was using the following cameras at the beginning of the war:

F-1	In-lens	5×7	8½" f:4.0 ("Altimar," Fig. 8)
			10" f:4.5
			20" f:5.6 (Telephoto)
F-8	Foc. Pl.	5×7	10" f:4.5
K-3, 3A, 3B	In-lens	18×24 cm	12" f:5.0
F-56	In-lens	7×7	40" f:8.0 (Telephoto)
			20" f:5.6 (Telephoto)
			8½" f:4.0 ("Altimar")

The F-56 series and the K-25 were used for fixed installation oblique photography. The K-20 and F-8 cameras (the latter equipped with a 15-inch f:5.6 Telephoto, see Fig. 7) were used for making hand-held oblique reconnaissance photographs. Later in the war the F-56 was replaced by the K-17 series and the K-18 for reconnaissance mapping because of the larger picture size. The K-19 camera was used for night photography.

THE STRIP CAMERA

This development was pioneered before the war by Colonel Goddard, and dozens of strip cameras have been manufactured and put into service. The principle is similar to that of the old rotating panoramic camera, in that the film is made to move forward continuously across a narrow slit in the focal plane of the lens at exactly the same speed as the image of the ground moves forward because of the motion of the airplane, no shutter being employed. In the later models such as those manufactured by F. Sonne of Chicago, the velocity of the film is controlled automatically to follow changes in the altitude and speed of the airplane. The Navy employed the strip camera mainly as a stereoscopic unit by using two 4-inch f:3.5 Ektar lenses (Fig. 2) side by side in front of a 9½-inch wide film, one lens being set slightly ahead of the other so that the continuous record of ground detail on one side of the film was taken a fraction of a second later than that on the other side. Because the airplane had moved a finite distance during the passage of this short interval of time, the two views were recorded from widely separated viewpoints, ideal for stereoscopic observation. Determining height of objects or depth of water by parallax measurement on parallel strip photographs is essentially similar to the pro-

TABLE I.

Camera	Shutter	Format	Lens	Remarks
K-17	In-lens	9×9	6" f:6.3 ("Metrogon")	
			12" f:5.0	
			24" f:6.0	
K-18	In-lens	9×18	24" f:6.0	
K-19	In-lens	9×9	12" f:2.5	Night-photo camera with photo-cell to operate shutter
K-20	In-lens	4×5	6½" f:4.8	Hand held
K-21	Foc. Pl.	5×7	7" f:2.5	For night photo
K-22	Foc. Pl.	9×9	(Same as K-17)	
K-24	Foc. Pl.	5×5	7" f:2.5	Night photo
			12" f:5.0	
			20" f:5.6 (Telephoto)	
K-25	In-lens	4×5	6½" f:4.8	An automatic form of K-20 camera

cedure involved in using ordinary stereo pairs. Prints from equivalent parts of the two strips may be cut and arranged under the stereoscope ("height finder") in the same way as an ordinary stereo pair, orienting each print so that measurement will be made in the direction of flight. Parallax may also be measured by a stereoscopic viewer having a parallax attachment which is so designed as to permit examination and measurement of the parallel strips without the necessity of cutting and arranging photos under an ordinary "height finder." Stereoscopic observations of this type, especially those in color, were of incalculable value to the Navy in all the attacks on the Pacific Islands, particularly for determining ocean depths near landing beaches. For the latter purpose, planes carrying strip cameras would fly at 400 miles an hour almost at tree-top height under conditions when the taking of a stereoscopic pair of single photographs would be virtually impossible with any other type of available camera.

Another method of using stereo strip photographs, which was devised but very little used, was to print the double film and run it through a twin projector with perpendicular polaroids over the lenses. The observer, of course, used Polaroid spectacles oriented in such a way that each eye would see its correct picture and ignore the other.

PROGRESS IN AERIAL LENS DESIGN

At the start of the war, little attention was paid to the ultimate ground resolution obtainable in aerial photographs, and indeed the various factors controlling this resolution were inadequately understood. The emphasis at that time was naturally placed mainly on production, but as the war progressed and immediate procurement requirements were met, it became increasingly and distressingly evident that the ultimate resolving power of most aerial photographs taken from high altitudes was quite inadequate to show small ground details with the clarity and sharpness necessary for military purposes. Many extensive investigations were therefore undertaken by the armed forces, the N.D.R.C., and by various manufacturing concerns to try to discover the underlying reasons for this poor performance and some way to rectify it.

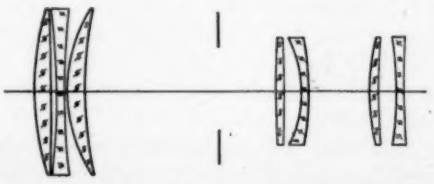

FIG. 9. Harvard University 40″ f:5 Telephoto.

FIG. 10. Bausch and Lomb 40″ f:8 Telephoto.

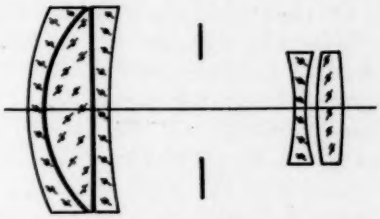

FIG. 11. Bausch and Lomb 40″ f:5.6 Telephoto.

This question had become prominent during World War I, but it was not solved then, and indeed we cannot honestly say that it is fully solved now. However, the intensive efforts of many workers along many different lines have made the problems very much more clear, and if the war had lasted a year or two longer it is certain that the definition obtainable in all aerial photographs would have become drastically improved.

One of the first fruits of these investigations was the development of a telephoto lens of novel construction (Fig. 9) designed by Dr. James G. Baker of the Harvard College Optical Laboratory for the N.D.R.C. It was a 40-inch f:5 lens equipped with temperature and pressure stabilization, to be used on the K-22 camera (9″×9″). Over 100 of these lenses were eventually manufactured, some at Harvard and some by Perkin and Elmer, and the ultimate resolving power attained with them exceeded 20 lines per millimeter, sometimes reaching as high as 40

lines per millimeter in actual aerial photographic tests.

When the war started, nobody liked the somewhat large K-18 camera with its 9″×18″ format, but by 1945 it was rapidly becoming the most popular size. However, it was soon found that the 24-inch f:6 lenses used with it had inadequate resolving power and were of too short a focal length, and consequently in 1944 the Air Corps invited designs for a 36-inch f:8 telephoto lens to fit directly into the shutter of the K-18 camera without further alteration except for any necessary small adjustment of the focusing studs. Samples of this 36-inch lens were submitted by two manufacturers and plans were being made for extensive production when the war ended.

Other new developments undertaken on a similar experimental basis were the Kodak 48-inch f:6.3 telephoto lens for the K-18 camera, and also the Harvard 60-inch f:6.3 telephoto and 100-inch f:10 quadruplet airspaced lens to be built into special folded cameras using 9″×18″ film. Prototype models of these huge cameras were approaching completion at the end of the war.

The Bausch and Lomb Company had manufactured a 40-inch f:8 telephoto lens for several years before the war. In 1940 this was redesigned (Fig. 10), and in addition a 40-inch f:5.6 telephoto (Fig. 11) was put into extensive production. It was a 5-element lens of more conventional construction than that designed at Harvard Observatory.

A small point worth recording is the universal elimination of distortion in the larger telephoto lenses developed during the war. This aberration was formerly considered inevitable in lenses of the telephoto type, but by airspacing the crown and flint elements in the negative component, it is possible to eliminate distortion almost completely.

THE FACTORS WHICH LIMIT THE RESOLVING POWER OF AN AERIAL CAMERA

There are many specific adverse factors which serve to limit the ultimate resolution obtainable in an aerial photograph. Although some of these factors have little direct bearing on the design of lenses, they are mentioned for the sake of completeness, and because they must be considered in planning lens cells and mountings.

The principal contributary causes of low resolution are:

(a) *Rolling and pitching* of the airplane. This is a mechanical problem which will undoubtedly be solved eventually, but today no satisfactory means are available for holding a camera on an "even keel."

(b) *Vibration.* Antivibration mountings are coming more and more into use, and it may well happen that vibration will eventually cease to be a contributing factor. The vibration caused by the opening and closing of a heavy camera shutter under powerful spring tension may at times cause a noticeable loss in definition.

(c) *Forward motion.* As the airplane travels along, the image of the ground moves forward across the focal plane of the lens at a rate depending on the height and speed of the plane and the focal length of the lens. This effect can be minimized by the use of a very short exposure time (1/100 or 1/150 second is generally employed), and recently some large shutters have been developed giving exposures as short as 1/500 second. However, the effect of forward motion can be eliminated entirely if the film is caused to move forward in the camera during the exposure at the same rate as the image is moving, and cameras embodying this feature were in course of development when the war ended. This procedure is of course followed continuously in the strip camera described above.

(d) *Atmospheric haze.* The effects of atmospheric haze become increasingly important as the altitude of the plane is increased. In the absence of definite white fog, the haze particles in the atmosphere lead to Rayleigh scattering which is proportional to the inverse fourth power of the wave-length. Thus, blue light is scattered strongly, yellow light less, and red light very little. Photographs taken with infra-red light on a very clear day show scarcely any effects of haze for object distances approaching 100 miles in some cases. In spite of this, infra-red film was not used very much during the war for reconnaissance purposes as it is rather slow, and the unfamiliar appearance of natural objects (especially trees) in infra-red light would necessitate a re-training of the personnel who have to

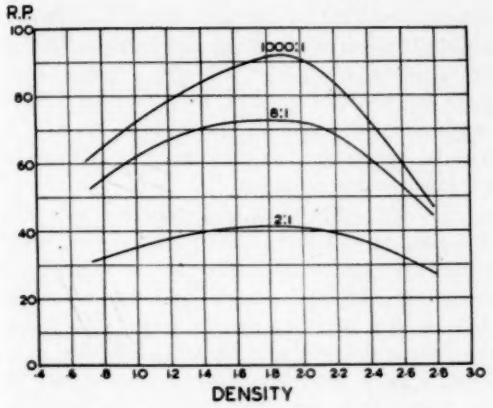

FIG. 12. Curves of resolving power against density for a typical aerial pan emulsion.

interpret the photographs. Deep red filters can be used with ordinary panchromatic film, but the necessary increase in exposure time is often undesirable. Most black-and-white photographs were therefore made with a minus-blue (yellow) filter and ordinary class L panchromatic film.

An increasing amount of aerial color film is now being used. Obviously no colored filter can be employed with this film, and the effect of atmospheric haze is to cover the picture with a blue coloration and to lower the contrast especially at the blue end of the spectrum. The film manufacturers do their best to help the situation by balancing the emulsion layers for an average amount of haze corresponding to, say, that at 20,000 feet, and by keeping the overall contrast as high as practical. The use of color film is increasing, and most future aerial camera lenses will require to be achromatized in the visible region of the spectrum rather than the former type of achromatism which ignored the blue end of the spectrum entirely. Even apochromatic correction is now being considered, as suitable crystals and new types of glass are becoming available.

(e) *The resolving power of the emulsion.* It is a most unfortunate fact that the resolving power of a photographic emulsion falls off rapidly from its optimum value, when the contrast in the object being photographed is low. The average contrast of an ordinary landscape containing possibly camouflaged objects is low, of the order of only 2:1, and at that contrast the resolving power of the emulsion may be only about half of its resolving power for an object of infinite contrast. The curves shown in Fig. 12 represent the variation in resolving power with density for a typical modern aero pan emulsion, the object contrast being indicated on each curve. It will be clear from this chart that instead of the optimum value of 92 lines per millimeter, this emulsion will yield a maximum resolution of only about 40 lines per millimeter under ordinary military conditions, assuming that the exposure time and development conditions are adjusted to give the optimum image density.

This condition can be improved radically only by changing to a low speed, high resolution emulsion such as a process or microfile emulsion. This would involve a many-times increase in exposure time, which is considered impractical at present because of the mechanical factors (a), (b), and (c) outlined above.

(f) *The choice of focal length.* It was very definitely established in the first World War that the longer the focal length of the lens the better was the ultimate definition of the picture, and to this day we do not appear to have reached the point of diminishing returns so far as long focal length is concerned. This is because the lens must form a sufficiently large-scale picture so that the fine detail we wish to record will be larger than the practical resolving power of the camera and emulsion, which is at present normally about 10 lines per millimeter under service conditions. However, it must not be forgotten that as the focal length of a lens is increased (by scaling up the lens formula to a larger size) the aberrations increase in the same proportion. Consequently the ultimate ground resolution obtainable with a given aerial camera when flown at a given height will be low when the focal length is short, and it will rise asymptotically to a fixed value for very long focal lengths, the final limit being set by the aberrations of the lens. It is thus of the greatest importance that the aberrations should be very well corrected in lenses of long focal length. At the start of the war the longest lens used by the Army was the 24-inch f:6.0, although the Navy was using the early design of 40-inch f:8 telephoto made by Bausch and Lomb. By the end of the war, as

TABLE II. Change in focus from ∞ setting.

Height of plane	$f=20$ inch	40 inch	60 inch
5000 ft	0.17 mm	0.68 mm	1.52 mm
10,000	0.09	0.34	0.76
20,000	0.04	0.17	0.38
30,000	0.03	0.11	0.25
40,000	0.02	0.09	0.19
50,000	0.02	0.07	0.15

has been mentioned above, contracts were being signed to make lenses of 60-inch and 100-inch focus, and even longer lenses were being considered. The chief limiting factor in the construction of huge lenses of this kind is the problem of procuring sufficiently large disks of homogeneous optical glass of the required type.

THE FOCUSING OF LARGE AERIAL LENSES

One very grave mechanical problem associated with the use of aerial lenses of long focal length is that of maintaining the image in exact focus upon the film at all times. Four factors affect the focus of the lens, namely, (a) a change in the color of a filter (because of the secondary spectrum of the lens) and any slight accidental lens power in the filter itself, (b) changes in object distance or height above the ground, (c) changes in ambient temperature, and (d) changes in atmospheric pressure due to flying height. These factors will be discussed in order.

(a) The change in focus of a 40-inch telephoto lens with normal photographic achromatism when going from a yellow filter to an infra-red filter may be as much as 4 or 5 millimeters, and even when yellow-red achromatism has been employed it may be over 1 millimeter. It is thus absolutely essential to re-focus the camera after such a filter change has been made. To avoid refocusing it has been frequently proposed to embody sufficient positive lens power in the glass of an infra-red filter to bring the image back into focus, but this would be successful only if each filter remained always associated with the particular lens for which it was intended. A similar proposal has been to insert the filters behind the lens and to adjust the thickness of each filter to maintain the lens accurately in focus.

(b) The change in focus due to object distance becomes decidedly significant when a lens of long focus is used, as indicated in Table II. There is no alternative to changing the focus setting of the camera when the altitude changes from 5000 to 40,000 feet if a long focused lens is employed, although for mechanical reasons, it may be desirable to adjust only one element of the lens for this purpose.

(c) If the *temperature* of a camera is suddenly changed, as in a swift ascent from a hot desert airport, very great internal strains are developed in the lens elements, and if the latter are six to eight inches in diameter by some two inches thick, as is the case in a 48-inch or 60-inch telephoto lens, it may take two to three hours for the temperature gradients within the lenses to die away. During the strain period, the image definition given by the lens will deteriorate badly, and this factor more than any other may be the cause of the early complaints that a good lens on the ground often made very poor photographs in the air.

However, quite apart from these transient effects caused by temporary temperature gradients in the lens, temperature changes cause a definite and readily measurable shift in the focal plane. This is owing to a combination of a physical change in the radii of curvature and thickness of the lens elements as a result of ordinary thermal expansion, a change in the lens separations because of the expansion of the barrel, a change in the refractive index of the air between the separate lens elements and in the camera body, thermal expansion of the camera cone, and a thermal change in the absolute refractive indices of the glasses themselves. The latter effect is actually very small at temperatures of 0° to −40°C, as that is the region in which most optical glasses change from a positive to a negative temperature coefficient of refractive index. In attempting to compute temperature effects in a lens, it is important to distinguish between the absolute index of the glasses (in vacuum) and the relative indices against air, because the index of air changes noticeably over the working range of atmospheric temperatures, which may run from +40° to −50°C for a high flying airplane in tropical climates. A theoretical study of these temperature effects in lenses has been published by J. W. Perry.[2]

[2] J. W. Perry, Proc. Phys. Soc. **55**, 257 (1943).

The following Table III has been taken from a recent paper by E. B. Woodford and R. N. Nierenberg[3] of Wright Field giving the measured shift in the flange-focus of several standard lenses for a temperature change of 1°C. For further details, the original paper should be consulted.

The focus shift can, of course, be compensated by re-focusing the camera according to the temperature of the lens, but this will not overcome the transient strain due to a temperature gradient. The only solution, therefore, appears to be that lenses and cameras must be fully thermostated so as to hold the initial ground temperature throughout the flight. This has been successfully accomplished by J. G. Baker in the Harvard 40-inch f:5 telephoto lens, and undoubtedly a similar device will be fitted to all large aerial lenses in the future.

(d) So far as *pressure* is concerned, it is easy to see that when a camera is exposed to the outer atmosphere, the density and consequently also the refractive index of the air inside it will change with altitude, quite apart from any temperature changes. This inevitably leads to a shift in the focal plane, and the values found by Woodford and Nierenberg for a series of lenses are included in Table III above. The pressure effect can be nullified by enclosing the lens in a sealed chamber held constantly at normal ground-level atmospheric pressure, but as this is generally impractical, some means must be provided to focus the lens for the height at which it is expected to be used. In Baker's 40-inch f:5 lens, the rear lens element is moved automatically by a device similar to an aneroid barometer in such a way as to maintain the image constantly in correct focus for any changes in the atmospheric pressure.

It is interesting to note that the temperature and pressure effects on a long-focused lens in the upper atmosphere are opposite in sign and decidedly significant in magnitude. For instance, taking Woodford and Nierenberg's data on the 40-inch f:8 lens in Table I, we find that during a flight from sea level to 40,000 feet, the atmospheric pressure will drop from perhaps 750 mm to 150 mm, which will lead to a shortening of the flange-focus of the lens by 0.8 mm. On the other hand, in such a flight the temperature may drop from +25°C to −55°C, a change in temperature of 80°C, leading to a lengthening of the flange-focus by about 1.5 mm. These values compensate each other to some extent, but it should be noted that the camera cone, if it is of aluminum, will shrink by perhaps 0.9 mm for an 80°C drop in temperature, which makes the whole situation decidedly worse and results in an error in focus of about +1.6 mm. By holding the temperature of the entire lens and camera constant, this focus error becomes that due to pressure alone, namely −0.8 millimeter.

TABLE III.

Lens	Flange focus shift per 1°C	Shift of focus per 1 mm change in atmospheric pressure
48-inch f:6.3 (Telephoto)	−.0195 mm	
40-inch f:8.0 (Telephoto)	−.0334	.00134 mm
Same lens in invar barrel	−.0117	
36-inch f:8.0 (Telephoto)	−.0294	
*24-inch f:6.0 (Tessar Type)	−.0214	.00049
*24-inch f:6.0 (Tessar Type)	−.0090	.00043
12-inch f:5.0 (Tessar Type)	−.0105	.00013

* Different designs by different manufacturers.

In preparing this paper, the author has received much useful help from Dr. W. B. Rayton, Dr. J. G. Baker, Mr. C. Ressler, and Mr. R. G. Tarkington. No mention has been made of a considerable number of special lenses which were developed experimentally during the war but for various reasons were never carried over into actual production.

[3] E. B. Woodford and R. N. Nierenberg, J. Opt. Soc. Am. **35**, 619 (1945).

Spectral and Modulation Characteristics of the Concentrated-Arc*

W. S. HUXFORD AND J. R. PLATT**
Department of Physics, Northwestern University, Evanston, Illinois
(Received October 24, 1946)

Measurements of the times of decay of light emitted in various spectral regions by the concentrated-arc indicate that the principal part of the visible and infra-red radiation is a continuum having its origin in a thin solid or liquid surface film on the arc cathode.

INTRODUCTION

A RECENT paper on "The Concentrated-Arc Lamp"[1] describes the static characteristics of this new light source developed by the Western Union Telegraph Company. The arc burns in a rare gas at atmospheric pressure between two electrodes spaced a few hundredths of an inch apart. The cathode consists of a tungsten or tantalum tube, or rod with a small cavity, filled with a zirconium oxide mixture. The chief source of intense light is the cathode spot formed on the oxide surface and having a diameter of 0.003 of an inch in the 2-watt lamp and 0.060 of an inch in the 100-watt lamp. The present report summarizes the results of a study of the dynamic characteristics of this arc and gives evidence to show that the intense continuous spectrum arises from the cathode itself and not from gases or vapor excited in the region of the cathode glow or in the arc column.

Photometric measurements show that the intensity of radiation in the visible spectral region ranges from 40 to 60 candlepower per mm² for the 100-watt concentrated-arc lamp. The intensity of light emitted by the anode of a normal carbon arc is approximately 160 candlepower per mm², which corresponds to the sublimation temperature of carbon, 3950°K.[2] In the case of the concentrated-arc, the radiation density corresponds approximately to that which would be obtained from a blackbody at 3200°K. The peak of the radiation curve of the 100-watt concentrated-arc, however, occurs at about the same wave-length as the peak of the curve of a tungsten filament operated at a brightness temperature of 2800°K. On either side of the peak the curve for the concentrated-arc lies somewhat above that for the tungsten filament. These considerations indicate that the continuum emitted by the concentrated-arc may not be caused entirely by near blackbody radiation from the heated cathode spot. The question arises as to how much of this radiation is excited in the gases and vapors in the immediate neighborhood of the cathode surface, and how much is owing to incandescence of the solid (or liquid) surface of the zirconium oxide cathode.

Many arcs are known to emit continuous spectra apparently originating at an electrode surface which, if interpreted as true thermal incandescence, would indicate very high electrode temperatures. There are theoretical reasons for believing that only a small part of this continuum is electron recombination spectra or "free-free" electron radiation. Buckingham and Deibert[1] observed in the vapor close to the cathode a weak continuum in the region from 3500A to 5000A which may be electron recombination bands. MacPherson[2] ascribed the intense continuous radiation from the anode crater of the high intensity carbon arc to transmission broadening of large numbers of spectral lines of elements used for impregnating the electrodes. Smith[3] has postulated a unique mechanism to account for the intense continuum in the negative glow of a mercury arc, and believes that it arises from cumulative electron excitation of the liquid.

* The work described in this paper was carried out, in whole or in part, under Contract OEMsr-990 between the Office of Scientific Research and Development and Northwestern University. The researches were completed in 1943–44 but publication was delayed due to restrictions imposed by military security. A first report of this work was presented at the Winter Meeting of the Optical Society of America, March 7–9, 1946, Cleveland, Ohio.

** Now at the University of Chicago, Chicago, Illinois.

[1] W. D. Buckingham and C. R. Deibert, J. Opt. Soc. Am. **36**, 245–250 (1946).

[2] H. G. MacPherson, J. App. Phys. **13**, 97–102 (1942).

[3] Charles G. Smith, Phys. Rev. **62**, 48 (1942); **64**, 40 (1943); **69**, 96 (1946).

The conclusion that the major part of the continuum observed in the concentrated-arc is excited in the surface layers of the cathode is based on measurements of the rate of decay of this radiation in the interrupted arc and in current modulated arcs. The decay constant associated with the modulated light has a value which is much too large to correspond to emission processes in the gas. A small fraction (5 to 10 percent) of the visible and infra-red radiation has a short period decay time and is ascribed to atomic lines or electron recombination spectra.

STUDY OF THE INTERRUPTED ARC

The current supplying a 100-watt concentrated-arc was interrupted periodically by short circuiting the arc with a commutator driven by a synchronous motor. The arc was interrupted once during each cycle of the 60-cycle a.c. for a period of 0.0028 sec., i.e. it was *out* one-sixth of each cycle, and *on* during the remaining five-sixths of the cycle. It was possible to operate some arcs in this manner and to produce restriking after each interruption only if a biasing potential in the form of a two- or four-volt storage battery was inserted in the leads from the arc terminals to the commutator brush. The result was a nearly perfect, square wave current pulse.

The rate of decay of radiation flux during the interruption period was observed by means of multiplier photo-tubes and a cathode-ray oscilloscope. For "blue" light, using an RCA 931 phototube, the instantaneous ($\tau < 10^{-5}$ sec.) decrease in light flux amounted to about 60 percent of the total, 30 percent had a more gradual decay, and about 10 percent was radiation having a very long period. For infra-red light detected by using Cs-Ag-O cathodes and Polaroid filters[4] which excluded all light having wave-lengths less than 8300A, the fraction ceasing instantaneously was approximately 20 percent, the decaying fraction 60 percent and the long period fraction about 20 percent. In both cases the time constant of the decaying fraction was of the order of $5 \cdot 10^{-4}$ sec. Only one early model of lamp was successfully operated in this type of test. It was found that a much more satisfactory method of studying time changes in the radiation was afforded by observing the light from an arc modulated with a sinusoidally varying current.

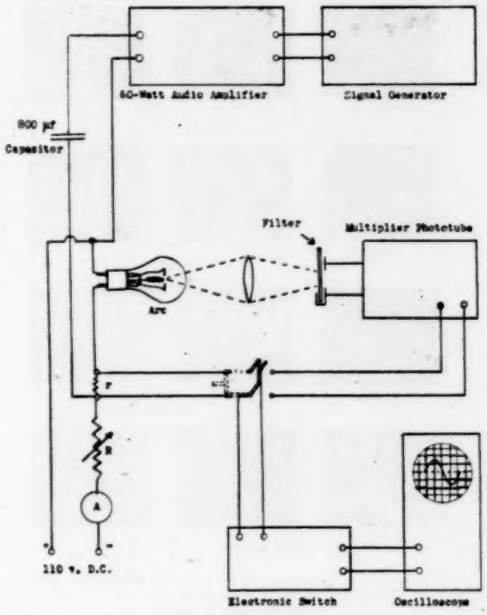

FIG. 1. Circuit for measuring light modulation ratios.

THE CURRENT MODULATED ARC

Figure 1 shows the arrangement used for modulating the concentrated-arc and for observing current, potential, and light waves by means of a cathode-ray oscilloscope. For current modulation at amplitudes having not more than 80 percent of the d.c. current value, a good sine wave form results if the power amplifier output impedance is properly matched to the arc impedance. An electronic switch was employed in conjunction with the oscilloscope for measurements of current and light modulation. Two types of multiplier photo-tubes were used, an RCA 931 for visible and ultraviolet light and a type 6PEA Farnsworth multiplier with Cs-Ag-O cathode for red and infra-red radiations. Filters were inserted in the light falling on the photocathode in order to isolate various spectral regions.

The dynamic characteristics shown in the plots of Fig. 2 are for 2-watt concentrated-arcs and are typical of all sizes of these lamps. The

[4] E. R. Blout, et al, J. Opt. Soc. Am. **36**, 460–464 (1946).

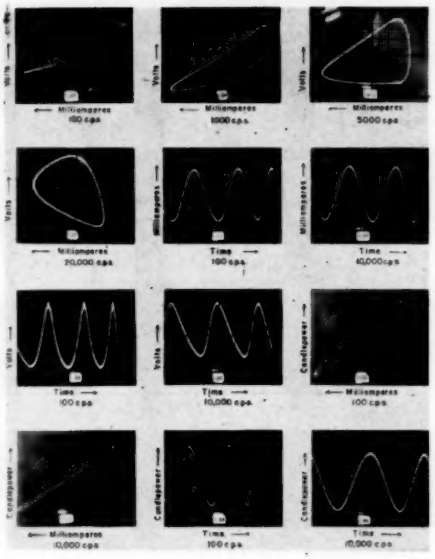

FIG. 2. Typical dynamic characteristics of the 2-watt concentrated-arc.

closed curves in the first four figures show the current-potential characteristic as a function of frequency of impressed alternating potential. (Current increases from right to left in these plots.) For low frequency of the order of 100 c.p.s. the alternating current follows nearly the d.c. negative resistance characteristic curve of the arc, with a 180-degree phase angle between current and potential. With increase in frequency the phase angle changes because of an inductive component introduced by the discharge, and also because the effective resistance of the arc changes from a negative value at low frequencies to a positive value at high frequencies.

The modulated light picked up by the multiplier photo-tube has essentially the same wave form as that of the current, and the waves are nearly in phase at all frequencies. The ratio of percent light modulation to percent current modulation, or the *light modulation ratio*, is nearly constant for current modulations up to values of about 80 percent. Some variations occur because the light received by the photo-cathode is a composite of gas spectra from the arc column and thermal radiations from the cathode surface. The results described here apply to tests carried out with fifty percent current modulation for which the wave forms are nearly sinusoidal.

MODULATION RATIOS USING WHITE LIGHT

Light modulation ratios were measured for a large number of arcs. Some arcs emit stronger line spectra than others, so that for measurements where white light is used, the ratios vary considerably for different lamps. The data shown in Fig. 3 were selected to illustrate this wide variation when light in the visible and ultraviolet portions of the spectrum is used. These results were obtained using an RCA 931 multiplier photo-tube and no filter. All curves show a characteristic double flexure with an approach to a horizontal section at low frequencies and at high frequencies. The tendency to a constant value at low frequencies indicates the presence of light which is not sensitive to rapid changes in current.

Figure 4 shows the results of measurements of modulation ratios in lamps operated at room temperature and then operated while immersed in a cold water bath. In most cases it was found that the light modulation ratios are higher for the cooled arcs. Presumably a component in the radiation due to general heating of the electrodes, which does not change rapidly with changes in arc current, is reduced in value when the lamp is cooled. The use of different filling gases does not seem to alter radically the effect of cooling on the light modulation ratios, as is shown in the results of Fig. 5. Not enough data were obtained with hydrogen filled lamps to determine

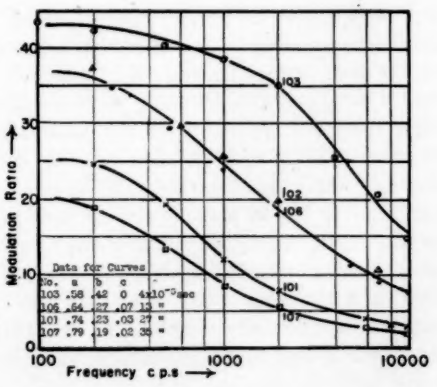

FIG. 3.

whether or not the modulation ratios are consistently higher when the arc is operated in this gas.

MODULATION RATIOS IN VARIOUS SPECTRAL REGIONS

It was found that modulation ratios for a given arc modulated at various frequencies decreased with increase in wave-length of the emitted radiation. This is illustrated for one arc in the results shown in Fig. 6. The points of the upper plot in this figure represent measurements using an RCA 931 tube and Corning filters 430 and 584, which pass light in the near ultraviolet. The data in the next plot were taken with the same photo-tube and a blue filter (Corning 430). For the points of the "red" plot, the Farnsworth photo-tube and Corning 240 filter were used. The measured points in the lower plot represent modulation ratios in the near infra-red using the Farnsworth photo-tube and Polaroid filter XR7X[4] which passes wave-lengths longer than 8300A. The infra-red threshold of this photo-cathode lies at about 12,000A.

Analysis of the results of these measurements indicated that for each arc, a part of the radiation is subject to a time lag with a period of the order of 10^{-4} sec. For a given arc the lag is apparently independent of the wave-length of radiation observed, and hence is a characteristic of the emission process. The details of this analysis are given below.

ANALYSIS OF MODULATION DATA

It is assumed that the radiation emitted at a given wave-length is made up of three parts:

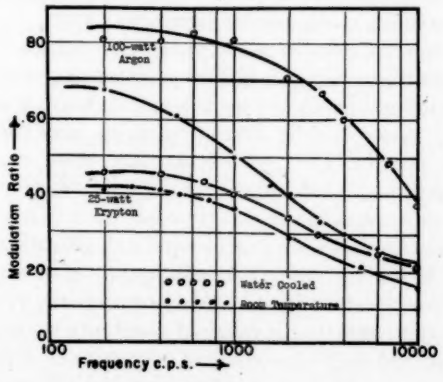

FIG. 4.

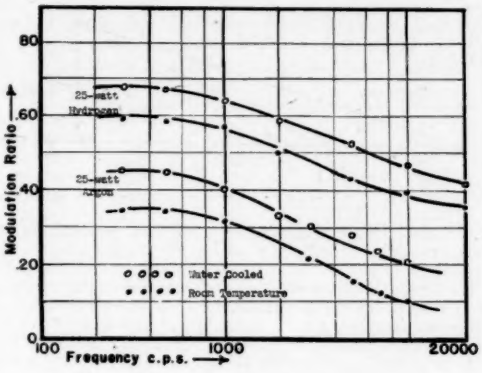

FIG. 5.

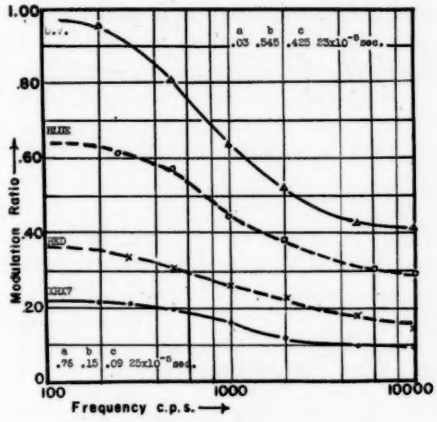

FIG. 6.

One part (thermal) susceptible only to long period fluctuations of average current, AI_{av}, where A is not necessarily constant with slow changes; one "modulatable" part with short-period exponential decay after excitation by current I, total effect given by sum of effects from all previous currents,

$$B\int_{-\infty}^{t} I(t')e^{-(t-t')/\tau}dt',$$

where τ is the time after current ceases for this part of the radiation to decay to $1/e$th of its previous value; and one essentially instantaneous part, CI, having a decay time much shorter than a single cycle of the exciting frequency. If the current varies sinusoidally $I=I_0(1+M\cos\omega t)$, where M is the percent current modulation, and

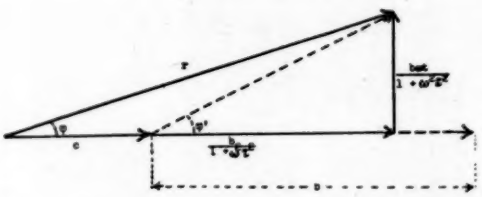

FIG. 7. Phase diagram for radiation emitted by the arc. r = light modulation ratio; $r^2 = c^2 + (b^2 + 2bc)\cos^2\varphi'$; φ = angle of lag; $\varphi' = \tan^{-1}\omega\tau$; b = component of the radiation having a decay time, τ; c = instantaneous component, in phase with the arc current; $a = 1 - b - c$ = unmodulatable thermal fraction, $t = \infty$. Values of c^2 and $(b+c)^2$ are obtained from plots of r^2 *vs.* $\cos^2\varphi'$, assuming values of τ. A check is obtained from experimental values of a, obtained at low modulation frequencies, since $(b+c) = 1 - a$.

the total light response will be given by

$$R = AI_0 + BI_0 \int_{-\infty}^{t} (1 + M\cos\omega t')e^{-(t-t')/\tau}dt' + CI_0(1 + M\cos\omega t). \quad (1)$$

Integrating (1) we obtain,

$$R = (A + B\tau + C)I_0 + I_0 M\left[B\tau \frac{1/\tau^2}{(1/\tau)^2 + \omega^2}\cos\omega t + B\tau \frac{\omega/\tau}{(1/\tau)^2 + \omega^2}\sin\omega t + C\cos\omega t\right]. \quad (2)$$

The *modulation ratio*, r, is obtained by dividing the amplitude of the second bracket of (2) by that of the first bracket. The "phase lag," ϕ, is just the phase lag of the second bracket of (2) behind $\cos\omega t$. Both of these quantities may be found from a phase diagram. For simplicity let

$$a = \frac{A}{A + B\tau + C}, \quad b = \frac{B\tau}{A + B\tau + C},$$

and

$$c = \frac{C}{A + B\tau + C},$$

a, b, and c thus being expressed as fractional parts of the total radiation at their wave-length, and $a + b + c = 1$. The phase diagram is shown in Fig. 7, where ϕ is the angle of lag of light wave behind current wave. In this diagram part c, with no lag, may be thought of as superimposed on part b having amplitude $b\cos\phi'$ and lagging behind c by the angle $\phi' = \tan^{-1}\omega\tau$. The light modulation ratio may now be expressed by

$$r = b\left(\frac{1}{1 + \omega^2\tau^2}\cos\omega t + \frac{\omega\tau}{1 + \omega^2\tau^2}\sin\omega t\right) + c\cos\omega t. \quad (3)$$

The qualitative predictions of the analysis based on these assumptions are easy to see. At low frequencies (small ω), ϕ' and ϕ are small and $r \to (c + b)$. At intermediate frequencies r decreases rather rapidly with increasing ω's and ϕ passes through a maximum at the frequency $\omega = (1 + b/c)^{\frac{1}{2}}(\tau^{-1})$. At high frequencies $\phi \to 0$, $\phi' \to \pi/2$ and $r \to c$. This behavior of r agrees with observed changes of the modulation ratio with frequency. Values of the phase angle ϕ were measured at the lower frequencies and found to increase with increasing ω's.

Fitting the theoretical curve to experimental values of r was made by trial and error to find the three constants b, c, and τ. In practice an easy way to obtain a fit is found by noting that

$$r^2 = c^2 + (b^2 + 2bc)\cos^2\phi', \quad (4)$$

and making plots of r^2 (function of ω) against $\cos^2\phi'$ (function of ω and τ) for several different values of τ. That value of τ is chosen for which the points make the best straight line. The intersections of this line with the lines $\phi' = 0$ and $\phi' = 90°$ are c^2 and $(b+c)^2$, respectively.

In Figs. 3 and 6 the values of b, c, and τ used in plotting the theoretical curves were found in this way. The points indicate experimental values of r, while the curves are plots of Eq. (3) using values of the parameters shown in the figures. Most of the r data taken on all models of arcs studied was represented by these formulas within the limits of experimental uncertainty.

For different arcs the values of τ varied considerably. (Fig. 3.) Values of τ for a given arc computed for curves for the various wave-length regions isolated by filters (Fig. 6) all agree within 20 percent. For a particular arc these curves could be fitted to the measured values of r within the range of experimental uncertainty by assuming a single value of τ and a single ratio b/c. It is realized that the assumption of a single decay constant is at best only a rough approximation, and that a range of constants would be expected for any statistical aggregate of similar processes. However, calculations based on the

use of various time constants show that the probable range of τ values could not be as great as 10:1. Hence we are reasonably certain as to the general nature of the mechanisms involved, since the mean decay time is of the order of 10^{-4} second, a value which is in agreement with observations of the decay of light in the interrupted arc.

Table I shows computed and measured values of the phase lag ϕ for the data given in Fig. 6. These measurements are difficult to make and may be subject to large errors at the higher frequencies. The predicted maximum in the value of ϕ was observed experimentally in only one case. The assumption of a spread of τ values leads to changes in the predicted phase lag ϕ. Qualitatively the lag angles at low and high frequencies are increased in value, and decreased at intermediate frequencies. The maximum changes are about 30 percent for a spread of τ over a range of 10 to 1.

TABLE I. Computed and measured values of the angle of lag, ϕ, of light wave behind current wave.

Frequency c.p.s.	Ultraviolet		Infra-red	
	Computed	Observed	Computed	Observed
250	12.5°	13°	10°	23°
500	22°	14°	17.5°	27°
1,000	28°	17°	23.1°	28°
2,000	24°	21°	21°	30°
5,000	13°	(31°)*	11°	(52°)*
10,000	7°	(42°)*	6°	(100°)*

* Measurements are subject to larger errors at the higher frequencies.

CONCLUSIONS

This analysis of the light modulation ratios indicates that the radiated energy represented by the unmodulated fraction a, which increases strongly with increasing wave-length, occurs at longer wave-lengths than the measured maximum of the total radiation. Presumably this radiation corresponds to a lower apparent temperature than the estimated 2800°K based on the radiation curve. Apparently this fraction is caused by general radiation from the heated electrodes which change their temperatures slowly with variations of arc current. It would be expected to vary greatly between types of lamps having widely different electrode structures, and was found to decrease when the arc was water cooled.

The fraction $(b+c)$ has a shorter wave-length maximum and hence a high apparent temperature, corresponding to, or greater than, the value of 3200°K estimated from radiation density calculations. Probably the fraction c is spectral radiation from the arc column or cathode glow, which in high pressure arcs will have a decay constant of the order of 10^{-7} second. The fraction b, which accounts for most of the modulated light in the infra-red region, has a decay time which cannot be ascribed to spectral radiations from gases or vapors in the arc. It is therefore to be associated with a mechanism which occurs in the surface of the activated zirconium oxide cathode. A surface temperature of 3000°K or larger is required to explain both the measured radiation intensities and the current densities observed in the concentrated-arc if thermionic emission of electrons occurs.

ACKNOWLEDGMENT

The writers wish to express their indebtedness to Messrs. W. D. Buckingham and C. R. Deibert of The Western Union Telegraph Company, Electronics Division, Water Mill, New York, for their valued aid in the design and development of many special types of arcs used in the course of these studies. This work could not have been carried out without their continued cooperation and the facilities of their laboratory. Mr. R. L. Osborn prepared the oscilloscope traces, and Mr. W. R. Wilson made the measurements of light-current phase lags.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 1 JANUARY, 1947

A Theory of the Anti-Coronae

H. C. VAN DE HULST
*Astronomical Observatory, Utrecht, Holland**

(Received October 14, 1946)

The anti-coronae, or luminous rings around the anti-solar point, are explained as a peculiar diffraction phenomenon in small water drops. The optical and the electromagnetic theories give results in agreement with each other and with the observations. The most important features are those which relate to the polarization.

THE diffraction coronae, or colored rings, which surround the sun or moon, when covered by a thin veil of cloud, are well known. A similar phenomenon can occasionally be observed in the opposite direction. A person standing on a high point observes his shadow projected on low clouds or on a layer of mist. He observes a gradual increase of the intensity of reflected light towards the shadow's head and, if conditions are favorable, some colored rings appear around the head. We shall call this phenomenon the anti-corona.[1] It is often seen in the mist on mountain tops and occasionally in lower country, when the sun is near the horizon. An air-pilot related that he saw the anti-corona around the shadow of his plane on the clouds nearly every day. According to the position of the observer, the center of the rings shifts from the shadow of the head of the plane to the shadow of its tail. Once or twice the anti-corona has been studied in laboratory experiments.

A satisfactory explanation has not so far been available. The old explanation by which the anti-corona is ascribed to common diffraction of light somehow reflected around the foremost drops of the cloud, does not appear to be sound. The more recent view is that these peculiar fluctuations in the intensity of nearly backward scattered light are already present in the scattering diagram of one single drop. Thus far, however, a more specific theory explaining the differences between the common coronae and the anti-coronae, has not been given.

For a general investigation of the scattering of light by spherical particles, we refer to a former publication,[2] which gives an extensive treatment of dielectrical and absorbing particles on the basis of Mie's electromagnetical theory. Data from other authors are compiled, various limiting cases are discussed and numerical results are given. The present problem deals with the limiting case of very large particles, where the scattering approaches the classical theories of geometrical optics and of optical interference. We shall first derive the results from the optical theory and then from the rigorous electromagnetical formulae.

OPTICAL THEORY

The intensity of light scattered by a sphere which is large in comparison with the wavelength can be computed by means of the laws of geometrical optics. The additional scattering by diffraction around the sphere is irrelevant for the present subject. The resulting formula is (T. 5, 14)

$$2\pi I_{1,2}(\theta) = \sum \epsilon_{1,2}{}^2 |a|^2. \qquad (1)$$

Here I is the emergent flux per unit solid angle, divided by the incident flux on the whole sphere; θ is the angle between the scattered and proceeding rays; finally the subscripts 1, 2 refer to the components the electric vectors of which are perpendicular and parallel, respectively, to the plane through the incident and scattered rays.

The intensities given by Eq. (1) consist of sums that have to be extended over all different light paths along which rays can be reflected, or refracted, into the direction θ. One light path can be specified by the parameters τ and p

* Now at the Yerlees Observatory, University of Chicago, Williams Bay, Wisconsin.

[1] The frequent use of the common term "glory" (cf. Webster) was considered unsuitable in an article of this kind.

[2] H. C. van de Hulst, "Optics of spherical particles," Thesis Utrecht, 1946; also published in Rech. Astr. de l'Observatoire d'Utrecht, 11, Part I. Formulae from this work will be quoted as (T. 5, 14) etc.

which are, respectively, the angle between the incident ray and the surface of the sphere, and the number of internal reflections added to 1. The case $p=0$ means external reflection. For a given light path, $\epsilon_{1,2}$ (T. 5, 13) denotes the decrease of the amplitude caused by the partial reflection and refraction on account of Fresnel's theory. Furthermore $|a|^2$ accounts for the decrease in intensity caused by the geometrical divergence of the rays; the value of $|a|$ is (T. 5, 15):

$$|a| = \left| \frac{\sin\tau\cos\tau}{\sin\theta} \frac{d\tau}{d\theta} \right|^{\frac{1}{2}}. \qquad (2)$$

It has been observed long ago that the intensity given by (1) and (2) would become ∞ for a light path for which

$$d\theta/d\tau = 0. \qquad (3)$$

This condition defines the *rainbows*. The two main rainbows lie near:

$$p=2, \quad \tau=31°, \quad \theta=138°, \quad \pi-\theta=42°;$$
$$p=3, \quad \tau=18°, \quad \theta=129°, \quad \pi-\theta=51°.$$

A strict infinity does not occur if interference is taken into account. The rays defined by (3) and the adjacent rays with equal p but with slightly different τ form a cubic wave front and a classical calculation[3] yields the correct intensity curve for the rainbows and supernumerary bows.

A second case in which the expression (2) diverges occurs when

$$\sin\theta = 0, \text{ while } \sin 2\tau \neq 0. \qquad (4)$$

No attention has been paid to this case, though it is quite similar to the case of the rainbow. Again, the divergence can be removed by taking into account the interference with the adjacent rays; a determinate intensity distribution will then result.

Condition (4) implies the cases $\theta=0$ and $\theta=\pi$. All peculiarities in the intensity distribution near $\theta=0$ are blended by the much stronger coronae caused by the diffraction of light around the sphere. Near $\theta=\pi$, however, a peculiar phenomenon may be expected. Indeed, as we shall see, the observed anti-coronae are explained in this way.

[3] Airy's theory. See any textbook on meteorological optics.

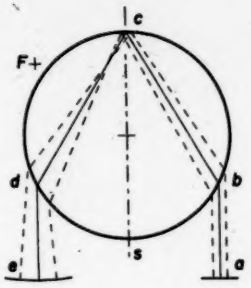

FIG. 1. Origin of toroidal wave front.

Figure 1 shows the simplest case of a light ray, *abcde*, satisfying the conditions required. Two adjacent rays, which emerge under slightly different angles, are also drawn. The linear front of the incident plane wave transforms into the circular front of the emergent wave, which has a virtual focus at *F*. However, still other rays must be taken into account. The whole figure must be rotated around the axis *cs*, and all the outgoing rays, thus defined, will interfere with each other. They define a *toroidal wave front* which seems to emerge from the focal circle described by *F*. We now have to calculate by means of Huygens' principle the interference pattern corresponding to this particular wave front.

The most interesting feature of the present problem is that the two directions of polarization cannot be treated separately. Let the incident wave be plane polarized with its electrical vector vibrating in the plane of Fig. 1. Then, for the rays drawn, it possesses a parallel vibration and emerges with an amplitude containing the factor ϵ_2. In the perpendicular plane, however, the same wave appears as a perpendicularly vibrating wave so that it is transmitted with an amplitude proportional to ϵ_1. Both transmitted waves are *still* vibrating in equal directions and are capable of interference. It is thus seen that the interference of rays in different azimuthal planes *simulates* an interference of rays with different directions of polarization.

Figure 2 explains the symbols we shall use in our analysis. The circle represents the focal circle from which the toroidal wave seems to emerge; let its radius be r'. We will compute the intensity radiated in a direction which makes a small angle γ with the axis and the projection of which on the plane of Fig. 2 is directed toward

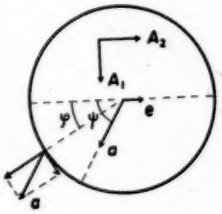

FIG. 2. Decomposition of light vectors.

the right, as shown by the arrow e. The consistent notations for the components of the light vector A of the emergent wave which vibrate parallel to and perpendicularly to the plane through the axis and e, are A_2 and A_1, respectively.

Let the incident light first be linearly polarized in a direction a making an angle ψ with the fixed direction mentioned. The light emerging from an arbitrary point of the focal circle, as specified by the angle φ, consists of two components which in the incident beam have the amplitudes $\cos(\psi-\varphi)$ in the outward radial direction and $\sin(\psi-\varphi)$ in the counterclockwise tangential direction. The amplitudes with which these components emerge are:

$$\begin{aligned} &\text{radial: } C_2\cos(\psi-\varphi)\\ &\text{tangential: } C_1\sin(\psi-\varphi)\end{aligned} \tag{5}$$

where the constants C_1 and C_2 are proportional to ϵ_1 and ϵ_2, respectively. To derive the amplitude vectors of the total emergent light wave, we must decompose the vector (5) into components with directions parallel to A_1 and A_2. These are:

$$\begin{aligned} a_1(\varphi)&=(C_1\cos^2\varphi+C_2\sin^2\varphi)\sin\psi-(C_1-C_2)\sin\varphi\cos\varphi\cos\psi,\\ a_2(\varphi)&=(C_1-C_2)\sin\varphi\cos\varphi\sin\psi-(C_1\sin^2\varphi+C_2\cos^2\varphi)\cos\psi.\end{aligned} \tag{6}$$

Writing $r'\sin\gamma=r'\gamma=u$, we now find the total amplitudes of the emergent light to be

$$A_{1,2}=\frac{1}{2\pi}\int_0^{2\pi}e^{-iu\cos\varphi}a_{1,2}(\varphi)d\varphi. \tag{7}$$

The definite integrals needed to reduce this expression,

$$\frac{1}{2\pi}\int_0^{2\pi}e^{-iu\cos\varphi}\cos^2\varphi d\varphi=\tfrac{1}{2}\{J_0(u)-J_2(u)\},$$

$$\frac{1}{2\pi}\int_0^{2\pi}e^{-iu\cos\varphi}\sin^2\varphi d\varphi=\tfrac{1}{2}\{J_0(u)+J_2(u)\},$$

$$\frac{1}{2\pi}\int_0^{2\pi}e^{-iu\cos\varphi}\sin\varphi\cos\varphi d\varphi=0,$$

are related to Sommerfeld's integral defining the Bessel functions.

Incident natural light can be considered as a superposition of non-coherent linearly polarized waves with randomly distributed directions of vibration. Averaging, therefore, the intensities in respect to ψ, we find a factor $\langle\sin^2\psi\rangle_{Av}=\langle\cos^2\psi\rangle_{Av}=\frac{1}{2}$ both in A_1^2 and in A_2^2. With omission of the factor $\frac{1}{8}$, the final intensities are:

$$\begin{aligned} I_1&=[C_1\{J_0(u)-J_2(u)\}+C_2\{J_0(u)+J_2(u)\}]^2\\ I_2&=[C_2\{J_0(u)-J_2(u)\}+C_1\{J_0(u)+J_2(u)\}]^2\end{aligned} \tag{8}$$

and in both directions of polarization together:

$$I=2(C_1+C_2)^2J_0^2(u)+2(C_1-C_2)^2J_2^2(u).$$

These formulae show the full implications of the interference effect described. If no interference between rays in different azimuthal planes took place, the intensities $I_{1,2}$ would contain only the coefficients $C_{1,2}$ with the *same* index. This is the case for the rainbow; it holds also for large angles in the present theory: with increasing distance from the center of the interference pattern the factor

$J_0(u)+J_2(u)$ decreases rapidly, like $u^{-\frac{1}{2}}$, leaving

$$J_0(u)-J_2(u)=(8/\pi u)^{\frac{1}{2}}\cos(u-\pi/4)$$

with the same index coefficient as the predominant term.[4] For small u, however, the terms of (8) are of the same order of magnitude; this region will be discussed in the last section.

ELECTROMAGNETICAL THEORY

The general expression for the intensity distribution of light scattered by a homogeneous sphere was first derived by Mie in 1908.[5] The resulting formulae in a somewhat changed notation (T. 2, 12) are:

$$\pi I_{1,2}(\theta)=\frac{1}{2x^2}|\sum_{1,2}|^2, \tag{9}$$

where

$$\begin{aligned}\sum_1&=\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\{a_n\pi_n(v)+b_n\tau_n(v)\},\\ \sum_2&=\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\{b_n\pi_n(v)+a_n\tau_n(v)\}.\end{aligned} \tag{10}$$

By x we denote the ratio $2\pi r/\lambda$, the circumference of the sphere in terms of the wave-length. The coefficients a_n and b_n refer to the electric and magnetic 2^n-poles, respectively; they are complex functions (T. 2, 7) of x and of the refractive index, m. Finally, the formula contains the spherical harmonics (T. 2, 10):

$$\pi_n(v)=\frac{1}{\sin\theta}P_n^1(v)\quad\text{and}\quad\tau_n(v)=\frac{d}{d\theta}P_n^1(v),$$

where $v=\cos\theta$.

Anti-coronae in visible light are observed on clouds consisting of drops with diameters of the order of 25μ, which corresponds to a value for x of about 150. For such a large value of x the evaluation of expressions (9) is impracticable so that we have to use their asymptotic forms. We denote the small angle $\pi-\theta$ by γ. The spherical harmonics have then the asymptotic forms (T. 5, 5):

$$\begin{aligned}\pi_n(v)&=(-1)^{n-1}\frac{n(n+1)}{2}\{J_0(z)+J_2(z)\},\\ \tau_n(v)&=(-1)^n\frac{n(n+1)}{2}\{J_0(z)-J_2(z)\},\end{aligned} \tag{11}$$

where we have written $z=n\gamma$.

As was stated by Debye, the terms of (10) are of comparable size up to $n=x$ and decrease rapidly as soon as n exceeds x. Most of these complex terms will nearly cancel each other. Only when the terms of the order

$$N-3,\ N-2,\ N-1,\ N,\ N+1,\ N+2,\ N+3,\ \cdots$$

have nearly equal phases, does an appreciable total amplitude result. In particular, the amplitudes of the nearly backward scattered light will be determined by the terms near the order N for which the sums

$$c_1=\sum_n(2n+1)(-1)^n b_n,\quad c_2=\sum_n(2n+1)(-1)^{n-1}a_n, \tag{12}$$

consist of terms with stationary phases. Neglecting the effect of the further terms, we find from

[4] By comparing this asymptotic behavior of (8) with Eq. (1) the relative coefficients may be expressed in absolute units.
[5] G. Mie, Ann. d. Physik 25, 377 (1908).

(10) and (11) the total amplitudes

$$\begin{aligned} \sum_1 &= \tfrac{1}{2}c_2\{J_0(u)+J_2(u)\}+\tfrac{1}{2}c_1\{J_0(u)-J_2(u)\}, \\ -\sum_2 &= \tfrac{1}{2}c_1\{J_0(u)+J_2(u)\}+\tfrac{1}{2}c_2\{J_0(u)-J_2(u)\}, \end{aligned} \tag{13}$$

where u is written for $N\gamma$. These formulae, when substituted into (9), agree with the optically derived formula (8). Only the values of c_1, c_2 and N have still to be calculated.

Before we proceed, attention may be drawn to the following point. In Debye's asymptotic formulae (T. 5, 8) the coefficients a_n and b_n refer to *one* direction of polarization. Yet they appear simultaneously in the amplitudes for *each* direction of polarization. This apparent paradox can now be completely explained. In the asymptotic form the term with the different index drops out for most directions because $\pi_n(v)$ has a smaller order of magnitude than $\tau_n(v)$. Only near the forward or backward directions the mixture of both terms remains and can be fully explained on the basis of Huygens' principle, as was shown in the preceding section.

The final step consists of the determination of N, c_1, and c_2 from (12). We may evaluate them either analytically or numerically.

a. By the analytical procedure we can formally demonstrate the complete equivalence of the optically and electromagnetically derived formulae for drops which are very large in comparison with the wave-length. Similar proofs have been given for the special case of the rainbow[6] and for the general case[7] in which expression (2) has no discontinuity. The modification for the present special case would present no difficulties. In each of the three cases we start by replacing the coefficients a_n and b_n by their asymptotic forms for very large order, according to Debye. Different expressions are needed for $n<x$ and for $n>x$, as can be visualized by means of the "localization principle," which assigns the terms of the order n to rays passing the center at a distance $n\lambda/2\pi$. Terms with $n>x$ correspond to rays passing along the sphere and give a vanishing contribution. Terms with $n<x$ correspond to rays hitting the surface under an angle $\tau=\arccos(n/x)$. Further, the asymptotic form for each term can be decomposed into subterms corresponding to different numbers of internal reflections, each of them with the proper coefficients ϵ_1, or ϵ_2. It appears that b_n refers to the perpendicularly vibrating wave (index 1) and a_n refers to the parallel wave (index 2). Finally, the subterms have a stationary phase for the order N which corresponds exactly with the location of the classical light paths. The sum of the terms of orders near N can be approximated by a Fresnel integral or, in the case of the rainbow, by an Airy integral. The resulting intensities, polarizations and phases agree with the optically derived results. For the anti-corona we would conclude that $N\lambda/2\pi$ is equal to the radius r' of the focal circle and that c_1 and c_2 are, like C_1 and C_2, proportional to ϵ_1 and ϵ_2.

b. The analytical derivation of N, c_1, and c_2 of which we gave an outline holds for the general case defined by Eqs. (4). However, our assumption that the actual anti-coronae are caused by light rays satisfying these conditions is not correct: a light path as illustrated in Fig. 1 does not exist for a waterdrop. When we pass from central rays to edge rays the deviation γ of the ray with a single internal reflection increases from 0° to 42° (rainbow) and decreases again to the final value of 14°. The value $\gamma=0°$ is not again reached. This would seem to invalidate our explanation. On the other hand, it must be pointed out that centrally incident rays, though they give $\gamma=0$, give a perfectly smooth intensity distribution since the discontinuity of expression (2) is removed by the fact that also $\cos\tau=0$.[8] Further, a rough computation shows that the more complicated light paths give rise to anti-coronae that are too weak to be observed, just as is the case for rainbows caused by more than two internal reflections.

The following simple solution is suggested as the most probable: though x can be as large as

[6] Balth. van der Pol and H. Bremmer, Phil. Mag. **24**, 141 and 825 (1937). Further articles on an analogous problem, *ibid.* **25**, 817 and **27**, 261.

[7] The outline given here follows the treatment by H. C. van de Hulst, reference 2, Chapter V.

[8] The explanation by B. Ray, Proc. Ind. Ass. for the Cultivation of Science **8** (1923), is not correct, for this reason.

200, the actual drops are still small enough to give strong deviations from the results for the optical case, $x \to \infty$. The following considerations tend to support this explanation. A light path as shown by Fig. 1 exists for refractive indices between 2 and $\sqrt{2}$. The angles of incidence required are given in Table I. It is obvious that $m = 1.33$ just fails to satisfy the conditions, if the optical theory is to be rigorously valid. For a finite value of x, however, the change at $m = 1.41$ will not be so abrupt. From the values of $\cos \tau$, which is equal to N/x, we see that the anti-corona is caused by rays near the edge of the droplet. We therefore infer that also for $m = 1.33$ the *radius of the focal ring is nearly equal to the radius of the droplet.*

Furthermore, we know that Debye's semi-convergent formulae for the cylindric functions, on which the proof of the equivalence between the optical and the electromagnetical results was based, are *not* valid for n near x. E.g., for $x = 150$ they are useless for the terms ranging from about $n = 144$ to 156.[9] This makes the optical theory for the anti-coronae completely unreliable for spheres with $x = 150$ and $m = 1.6$ to 1.41. Consequently, it is probable that drops of this size and of refractive index 1.33 still give a strong anti-corona.

The values of c_1 and c_2 cannot now be derived from the analytical theory. Instead, we should compute the values of a_n and b_n from their rigorous definitions and add the sums (12) *numerically*. The optical considerations only predict a stationary phase of the terms near $n = x$. We have not made such a computation and, accordingly, we cannot give a theoretical prediction about the ratio of c_1 to c_2.

TABLE I. Rays that would give rise to anti-coronae if the refractive index were m.

m	τ	$\cos \tau$
2.0	90°	0
1.9	54°	0.59
1.8	38°	0.78
1.7	26°	0.90
1.6	16°	0.96
1.5	7°	0.99
1.41	0°	1.00
1.33	—	—

[9] See Jahnke-Emde, *Tables of Functions* (B. G. Teubner, Leipzig, 1933 edition), especially Fig. 105.

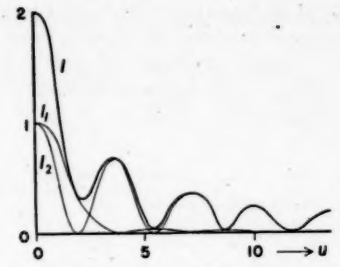

FIG. 3. Intensity distribution of the anti-coronae in both directions of polarization, in case $C_1 = 0$.

COMPARISON WITH THE OBSERVATIONS

For any given ratio of C_1 to C_2 the relative intensities in both directions of polarization can be readily evaluated from (8). The results for a few cases are:

1. $C_1 = C_2$. The anti-corona is wholly unpolarized. The central field is very luminous and dark rings appear at $u = 2.5, 5.6, 8.7, 11.8, \cdots$.

2. $C_1 = -C_2$. The anti-corona is again unpolarized. The anti-solar point is dark and is surrounded by a luminous ring at $u = 3.1$. Dark rings are situated at $u = 5.2, 8.5, 11.6, \cdots$.

3. $C_1 = 0$. The intensities for this case are shown by Fig. 3. In the central field the "alien" polarization, index 1, is slightly preponderant except at the anti-solar point itself. At $u = 2.3$ we find a fairly dark ring in which the polarization changes its sign. The bright ring at $u = 3.5$ and all further rings are nearly completely polarized in the "proper" direction, index 2. They are separated by dark rings at $u = 5.4, 8.6, 11.7, \cdots$.

We shall now compare these predictions with the available *observations*. Data concerning anti-coronae on natural clouds have been collected by Pernter-Exner.[10] Measurements on artificial mist, together with qualitative observations of the polarization, have been published by Mierdel.[11] These data, though not obtained with modern observational technique, suffice to give some checks on the theory and a preliminary determination of C_1/C_2.

The anti-coronae show distinct *differences from common coronae*, all of which can be explained on the basis of the present theory. (a) One feature is

[10] Pernter-Exner, *Meteorologische Optik* (W. Braumüller, Wien, 1910), p. 413 ff.

[11] F. Mierdel, Beiträge z. Physik d. freien Atmosphäre **8**, 95 (1919).

TABLE II. Ratios of the radii of the dark rings.

Observer	Light	Cloud	γ_1/γ_2	γ_3/γ_2
Mierdel	white	artificial	0.34±0.05	1.66±0.03
	red	artificial	—	1.68±0.04
	photographic	artificial	—	1.59
Average of several observations	white	natural	0.46±0.05	1.67±0.06
Wegener	photographic	natural	0.41	(1.74)
Theory	$C_1=0$		0.43	1.60
	$C_1/C_2=-0.25$		0.35	1.61

their variability. It is obvious that the interference of rays which are refracted by opposite sides of a waterdrop will be much more sensitive to slight deformations of the droplets than are the rainbows in which only adjacent rays interfere, or the common coronae in which non-refracted rays interfere. (b) The outer rings of anti-coronae are much more pronounced than are the outer rings of common coronae. On one occasion as much as 5 minima could be observed. The explanation is that the intensity in the anti-coronae decreases proportionally to γ^{-1}, whereas common coronae follow a θ^{-2}-law. (c) A final striking feature is the haziness of the first dark ring (which for normal sunlight is observed as the first red ring). Mierdel even calls it a slight depression separating the inner and outer parts of the central field. This haziness can be understood if the intensity distribution of Fig. 3 is approximately correct. A further increase of the first bright ring is obtained if we approach case 2, where C_1 and C_2 have opposite signs. Estimating from Mierdel's description that

$$0.30<\frac{\text{brightness of first ring}}{\text{brightness of central field}}<0.80,$$

we find the value of either C_2/C_1 or C_1/C_2 to lie between 0 and -0.25.

Further information is obtained from Mierdel's observations of the *polarization*. From the fact that the picture rotates with the analyzing nicol we infer circular symmetry to exist so that the glass plates in Mierdel's experiment cannot have had a serious effect. The complementary colors of the central field and the rings indicate different planes of polarization, in agreement with Fig. 3. The preponderant polarization in all bright rings is the one in which the electric vector vibrates in the radial direction. This shows that $C_2>C_1$, disregarding signs. Still, the tangential component is also visible. Estimating its intensity to exceed 4 percent of the total intensity of the rings, we find $C_1^2/C_2^2>0.04$. Together with the former data this indicates that C_1/C_2 must be about $-\frac{1}{4}$ or $-\frac{1}{3}$. This estimate is based on laboratory data; no observations of the polarization of natural anti-coronae seem to have been published.[12]

Finally we use the *radii* of the rings. As is true for common coronae, the radii of the rings are inversely proportional to the sizes of the droplets. However, it was noticed long ago that the rings of anti-coronae do not give consistent sizes when interpreted with the older theory. Let us denote the radii of the dark rings by $\gamma_1, \gamma_2, \gamma_3$, etc. For $C_1=0$, the present theory gives

$$\gamma_1:\gamma_2:\gamma_3:\gamma_4=2.3:5.4:8.6:11.7$$
$$=0.43:1.00:1.60:2.17.$$

For other values of C_1/C_2 the ratios are not very different. Table II shows the observed and computed ratios.

The agreement between theory and observations is as good as can be expected. No correction has been applied for the size of the sun and it was assumed that the edges of the red rings, which were measured by the observers, indicate the positions of the *dark* rings in the monochromatic picture for yellow light.[13] This introduces a considerable uncertainty but the ratio $\theta_2/\theta_1=1.80$, valid for the common diffraction coronae, differs strongly enough from the tabulated values to indicate again that this theory does not apply to the anti-corona.

On the basis of the present theory, measurements of the various dark rings give consistent sizes for the droplets. From the average diameters, which are about $2\gamma_1=1°\,30'$, $2\gamma_2=3°\,50'$, $2\gamma_3=6°\,30'$, we find 0.027 mm for the average diameter of the water drops in the clouds on which anti-coronae have been observed.

[12] Dr. O. Struve has related to me that he observed an anti-corona through a nicol prism on a transatlantic flight, Sept. 15, 1946. It was strongly polarized and the dark and bright features rotated when the nicol was rotated. No radii or intensities were however estimated.

[13] Early experiments by Fraunhofer indicated that this rule was correct for the first three rings of common coronae. The "wave-length of white light" is 0.57μ. See Pernter-Exner, p. 460.

Comparison of LiF and CaF_2 Prisms for Infra-Red Use

ROBERT C. GORE, ROBERT S. MCDONALD, AND VAN ZANDT WILLIAMS
Stamford Research Laboratories, American Cyanamid Company, Stamford, Connecticut

AND

JOHN U. WHITE
The Perkin-Elmer Corporation, Glenbrook, Connecticut

(Received September 4, 1946)

A PREVIOUS paper[1] discussed typical infra-red spectra and resolution obtainable from a Perkin-Elmer infra-red spectrometer. In that paper the results from LiF, NaCl, and KBr prisms were shown as obtained with both single and double galvanometer signal amplification and a galvanometer light beam follower. Since that time electronic amplification and recorders have replaced the galvanometer systems with consequent improvement in spectrum reproducibility and over-all instrument stability. Moreover, synthetic CaF_2 prisms have become available and the HF impurity absorption in synthetic LiF has been eliminated.

In view of the increasing need for better dispersion and resolution in the 2.5–7μ region a definite question arises as to relative performance and desirability of CaF_2 and LiF for this purpose. Therefore, the purpose of this paper is to present comparative results obtained from these crystals* on both low molecular weight gases and typical materials encountered in an industrial laboratory. For purposes of convenience and comparison with the previous paper some recent NaCl and KBr data are also included.

EXPERIMENTAL

The measurements were made with a Perkin-Elmer Infra-red Spectrometer Model 12-B, a General Motors breaker type d.c. amplifier, and a Brown Electronik recorder. This equipment was set up as normally used in a laboratory that was fairly free from electrical interference but was not air conditioned. The spectrometer was adjusted for best focus with each prism before measurements were made. The tubes in the amplifier were likewise tested for noise. The prisms have 60-mm height, 65-mm altitude and apex angles of 60° for NaCl, 69° for CaF_2, and 72° for LiF.

[1] R. B. Barnes, R. S. McDonald, V. Z. Williams and R. F. Kinnaird, J. App. Phys. 16, 77 (1945).

* The prism blanks for this purpose were kindly supplied us by the Harshaw Chemical Company of Cleveland, Ohio, and polished by The Perkin-Elmer Corporation of Glenbrook, Connecticut.

RESULTS

Comparison of Prisms for Resolving Power

Measurements were made by putting a sample in the light path and obtaining automatic records of the absorption spectra. The actual resolving power was determined by studying a number of compounds in appropriate regions of the spectrum and choosing pairs of bands or lines that were just resolved (i.e., showed an absorption

TABLE I. Resolving power of prisms.

Compound	Lines resolved	Separation	Rayleigh prism resolution	Ratio
		KBr		
CH_3OH	431 – 416 cm^{-1}	15 cm^{-1}	0.3 cm^{-1}	0.024
CH_3OH	506 – 501	5	0.56	0.11
CH_3OH	618 – 612	6	1.0	0.16
NH_3	854 – 848	6	2.1	0.35
NH_3	1096.1–1104.1	8	4	0.5
		NaCl		
NH_3	921.5– 918.9	2.6	0.85	0.33
H_2O	1424.1–1420	4.1	2.1	0.51
H_2O	1521.7–1517.1	4.6	2.4	0.52
H_2O	1675.5–1669.8	5.7	3.0	0.53
H_2O	1800.5–1793.3	7.2	3.4	0.47
C_2H_2	2963.2–2980.9	17.7	7.4	0.42
H_2O	3648 –3675	27	9.7	0.36
		CaF_2		
NH_3	1123.0–1117.4	5.6	0.3	0.055
CH_4	1262.6–1261.5	1.1	0.5	0.45
H_2O	1473.5–1472.1	1.4	0.58	0.41
H_2O	1603.6–1601.7	1.9	0.72	0.38
H_2O	1811.2–1808.7	2.5	0.91	0.36
CO	2201.1–2197.7	3.4	1.4	0.41
CH_4	3158 –3149	9	2.8	0.31
H_2O	3745 –3733	12	3.9	0.32
		LiF		
H_2O	1675.5–1669.8	5.7	0.3	0.053
H_2O	1773.2–1768.9	4.3	0.35	0.081
CO_2	2345.1–2343.4	1.7	0.7	0.41
C_2H_4	3100.7–3095.4	5.3	1.4	0.26
H_2O	3903 –3894	9	2.1	0.23

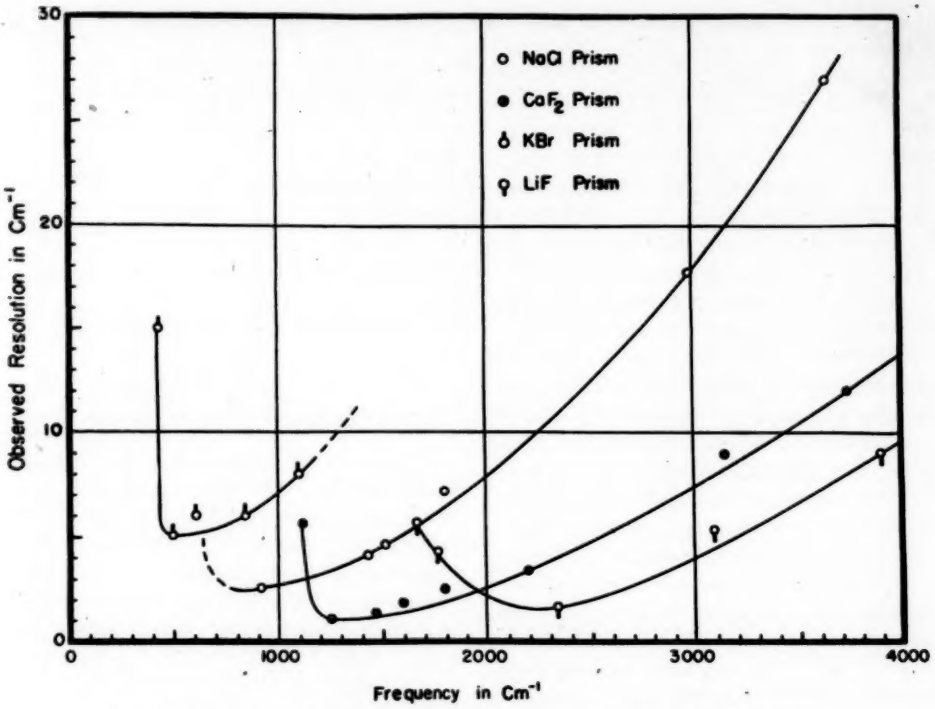

FIG. 1. Plot of observed spectral resolution *vs.* frequency for various prism materials.

minimum). The resolution measured in this way is a minimum which might have appeared higher if more closely spaced or sharper absorption lines had been available. The slit widths used for determinations were such as to give an I_0 energy signal of about a quarter of a microvolt or greater (signal to noise ratio of about one hundred). Frequency values for the lines studied were taken from grating measurements in the literature.

The pairs of lines used as criteria are shown in Table I along with the separation, the prism resolving power at that frequency according to the Rayleigh criterion, and the ratio of the observed to theoretical resolving power. These data have also been plotted as graphs in Fig. 1, where the resolution of each prism is shown as a function of frequency. It shows that the resolving power of each prism increases steadily and fairly uniformly with decreasing frequency until it reaches a maximum and suddenly decreases. The steady increase in resolution with decreasing frequency is caused by increasing dispersion and would continue to increase considerably further if absorption in the prism did not reduce the amount of light transmitted by it, thereby necessitating wider slits and lower resolution. It was actually found that the rotational structure of the 14.9-micron CO_2 band was resolved with a rocksalt prism if the slits were closed down far enough to give 20 percent deflection. The 1.9-wave number separation of these lines was not included in the tabulation or figures because it was not obtained with sufficient energy.

The curves show that there is an optimum prism for each part of the spectrum and indicate the relative performance of the different prisms at each point. A few of the conclusions which can be drawn from them are:

Of the materials considered the optimum prism for resolution in the different parts of the spectral range from 4000 cm^{-1} (2.5μ) to 400 cm^{-1} (25μ) is

Range		
Frequency (cm^{-1})	Wave-length (μ)	Prism
Less than 650	Longer than 15.4	KBr
650–1180	15.4–8.5	NaCl
1180–1900	8.5–5.3	CaF_2
More than 1900	Shorter than 5.3	LiF

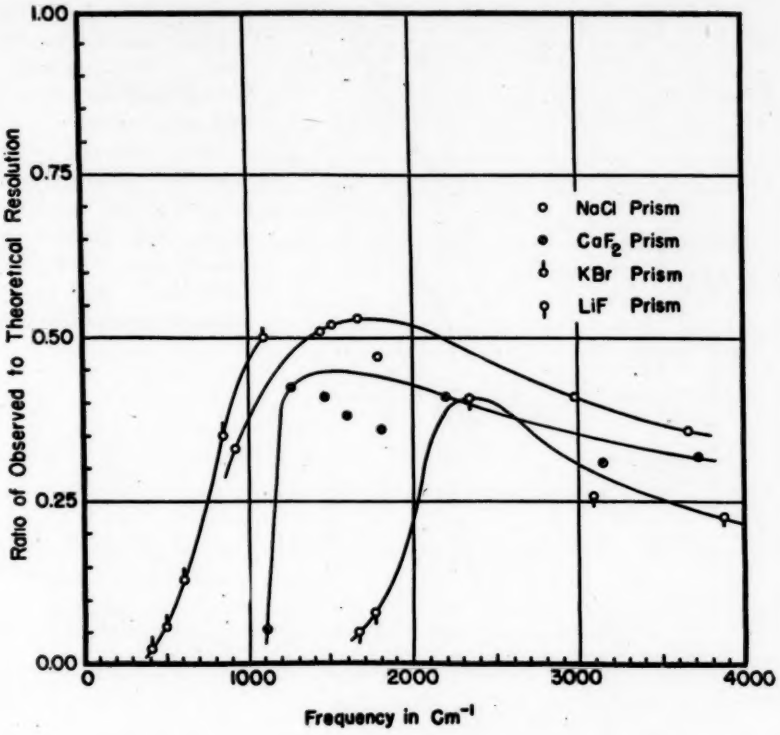

FIG. 2. Ratio of observed to theoretical (Rayleigh) resolution for various prism materials.

If a salt and LiF prism are considered, the point of equal resolution is 1660 cm^{-1} or 6μ.

In the 3-micron region the separations resolvable by three of the prisms are

LiF— 5 cm^{-1},
CaF_2— 8 cm^{-1},
NaCl—18 cm^{-1}.

This indicates that an improvement factor of about 2.2 is obtained in changing from rocksalt to fluorite, but that the subsequent improvement in changing from fluorite to LiF is only about 1.6.

In Fig. 2 the ratio of observed to theoretical (Rayleigh) resolution is plotted as a function of frequency for the different prisms. The curves show that the spectrometer's present performance approaches fairly close to the limiting possible performance in the optimum range in each prism. The loss in resolution at the left-hand end is caused by the reduction in light intensity by the prism absorption mentioned above. The decrease in resolution at the right-hand end of the curve is caused by loss in energy attributed to decreasing dispersion and consequently narrower slits. At very narrow slits, resolution is also reduced by the residual aberrations in the monochromator. This point is only just approached in the rocksalt curve. Measurements of the mercury yellow lines and sodium *D* lines show that the mercury lines can be well separated by the rocksalt prism and that the sodium lines can just be resolved with a glass prism and thermocouple detector.

Comparison of Prisms for Spectroscopic Applications

Most of the users of infra-red spectroscopy are not concerned with the ultimate resolution of the rotational structure in simple gaseous molecules such as hydrogen chloride, carbon dioxide, or ethylene, but are interested in the study of the structure and the quantitative analysis of complex molecules in the liquid, solid, and solution phases.

In such studies the rotational structure is unobservable and the vibrational bands are

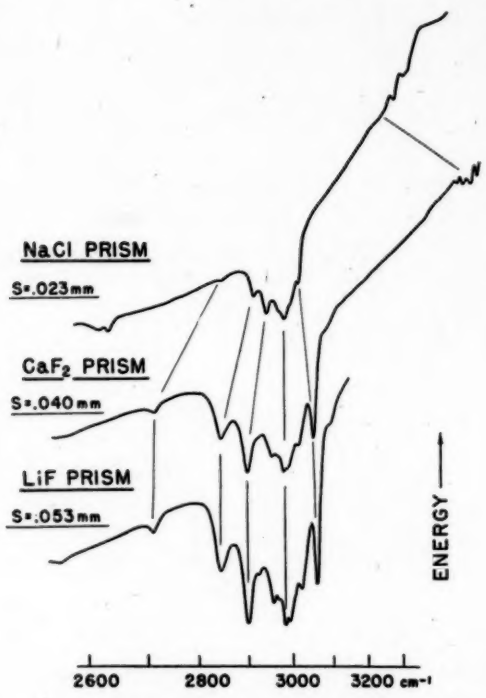

FIG. 3. The hydrogen stretching vibrations of *p*-methyl styrene.

usually of a width greater than 5 cm^{-1}. Increasing the resolution by using the prism with optimum characteristics in a given spectral region often adds no useful information. Accordingly, the choice of a prism material depends upon several factors such as: (1) the width and separation of the useful absorption bands in the studied compounds in the desired states; (2) the presence of an overlaying resolvable atmospheric absorption which may bring confusion in spectral interpretations; (3) the useful spectral interval encompassed by the prism material; (4) the angular dispersion of the prism which governs the mechanical accuracy of calibration; and (5) the cost and permanence of the prism material.

The exploitation of the frequency differences, correlatable with structure, that are known to exist[2,3] in the spectra of many liquids and solids in the hydrogen radial vibration and $X=Y$ stretching frequency regions requires resolution greater than that afforded by rocksalt. As a typical example the partial spectrum of *p*-methyl styrene recorded with rocksalt, fluorite, and lithium fluoride prisms is shown in some of the following figures.

In the 3000 cm^{-1} region, Fig. 3, the additional resolution afforded by CaF_2 and LiF prisms is illustrated for *p*-methyl styrene in the liquid phase. This enables the presence of aliphatic unsaturation (3060 cm^{-1}), the phenyl radical (3000 cm^{-1}, 3020 cm^{-1}) and the methyl group attached to a carbon atom (2840 and 2970 cm^{-1})

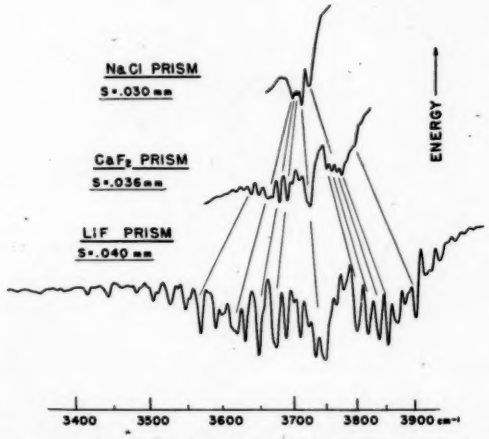

FIG. 4. Water vapor absorption in the 3μ region. The LiF prism is now free of the impurity band.

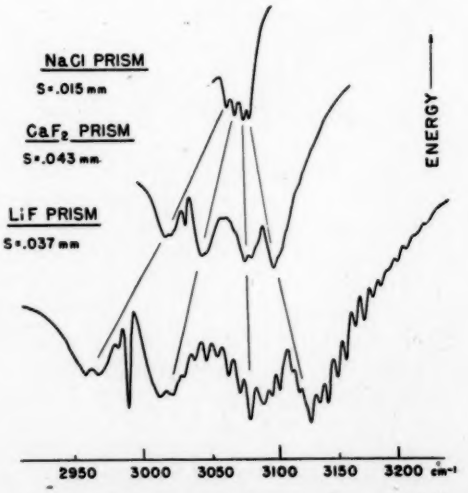

FIG. 5. The hydrogen stretching vibrations of ethylene.

[2] Fox and Martin, Proc. Roy. Soc. **A175**, 208–233 (1940).
[3] N. Wright, Rev. Sci. Inst. **15**, 22–27 (1944).

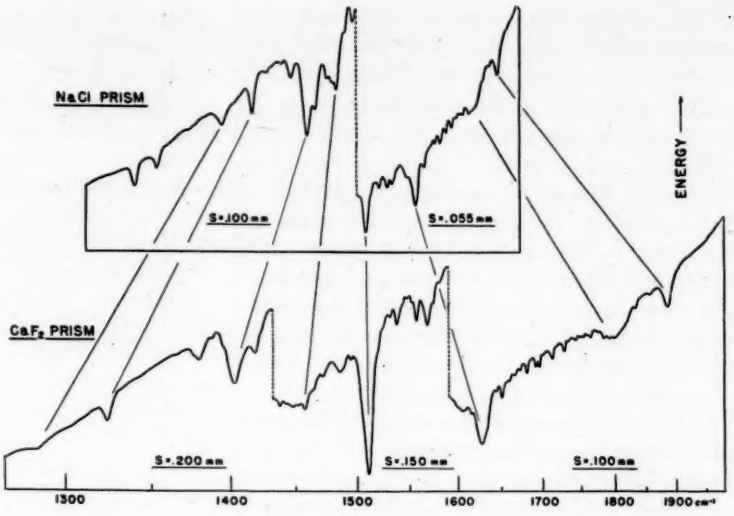

FIG. 6. Absorption of *p*-methyl styrene. Much more careful efforts were made to get rid of water vapor for the CaF_2 than for the NaCl spectra.

to be definitely established. It should be observed that for this compound the additional resolution afforded by LiF does not seem particularly more useful than that of the fluorite prism although the advantage of LiF in this region is obvious for light molecules as water vapor, Fig. 4, and ethylene, Fig. 5.

In the 1500 cm^{-1} region, Fig. 6, only rocksalt and fluorite prism spectra are shown because the energy obtainable through the lithium fluoride

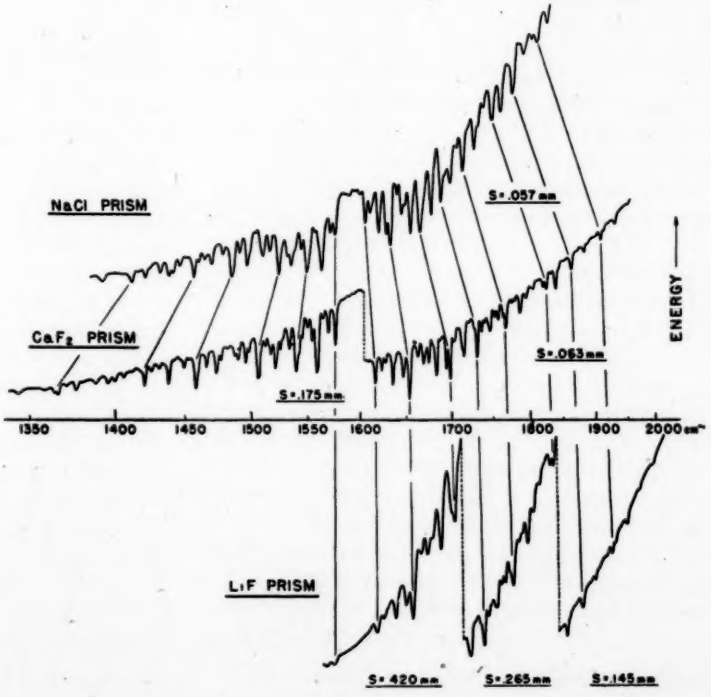

FIG. 7. The 6μ water vapor band.

prism at reasonable slit widths is too small and changes too sharply to be usable. The advantage in resolution afforded by the fluorite prism is evident especially if the absorption bands are being used in quantitative analysis. The detection and measurement of another component having an absorption band, say at 1650 cm^{-1}, would be more readily accomplished with the CaF_2 prism. The nature of the absorption bands of the *p*-methyl styrene, however, appears to be such that there is little further qualitative resolution with the fluorite although the actual increase of fluorite resolution is shown in the water vapor absorption of this region in Fig. 7.

From a practical standpoint then the following comparisons may be roughly summarized for the three prism materials.

Rocksalt: It has general usefulness from 4000 to 650 cm^{-1}. The resolution near 3000 cm^{-1} is insufficient for many applications and because of its low angular dispersion the accuracy of its calibration is not high in this region. Atmospheric absorption causes no great difficulties at rocksalt resolution.

Lithium fluoride: The useful region extends from 4000 cm^{-1} to 2200 cm^{-1}. Recent samples of synthetic lithium fluoride from the Harshaw Chemical Company do not show the troublesome F-H-F absorption peak near 3562 cm^{-1}. (Fig. 6.) If LiF is used at frequencies lower than 2200 cm^{-1} the rapid energy loss as shown in Fig. 7, necessitating frequent opening of the slits, gives a distinct disadvantage over the use of salt and fluorite. The low water solubility of the material makes for permanence in the optical finish of the prism.

Calcium fluoride: The useful region, with only slightly less resolution than that afforded by lithium fluoride, extends from 4000 to 1300 cm^{-1}. Use beyond 1300 cm^{-1} again has the disadvantage of a rapidly changing energy slope. Strong, highly resolved, atmospheric water vapor bands near 1600 cm^{-1}, while affording accurate calibration in this region, must be eliminated in order to obtain useful, unconfusing spectra. Fluorite offers no great advantage over rocksalt in the 1600 cm^{-1}, $X = Y$ region, however, unless a special study requires a maximum of resolution or of energy at a given spectral slit width. Fluorite's insolubility in water is even more than that of LiF.

If a choice must be made between fluorite and lithium fluoride for a prism auxiliary to rocksalt, in spite of its slightly lower resolution and because of its greater range the calcium fluoride would probably be desirable.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 1 JANUARY, 1947

Use of Incompletely Consumed Samples to Illustrate Effect of Fractional Distillation in Carbon Arc Spectrochemical Analyses*

LESTER W. STROCK AND GEORGE E. HEGGEN
Spectrochemical Laboratory, Saratoga Springs Commission, Saratoga Springs, New York

COMPLETE CONSUMPTION PROCEDURE

IN a method devised to determine the elements Mg, Ca, Sr, and Ba in saline brine residues, BeO was introduced as an intensity standard. This standard substance was chosen because of its chemical similarity to the elements to be determined. Moreover, the Be spectrum is very simple, and a line at 4572.671 lies in a favorable position with reference to the analysis lines employed in the visible spectral range. It was also expected that the presence of a large amount of BeO added to a NaCl base would increase the arc temperature and favor the volatilization of the other alkaline earth elements into the arc gas, particularly when admixed with carbon powder to reduce fractional distillation. The BeO intensity standard was introduced by mixing *one part of sample* (or standard) *to one part of an intensity standard: buffer mixture* containing 50 percent BeO in NaCl. Accordingly the mixtures *as arced* contained *25 percent BeO, and about 70 percent NaCl* since the samples themselves carried approximately 90 percent NaCl. These mixtures were finally combined with two parts carbon powder and packed into carbon electrodes ($\frac{1}{16}''\times 4$ mm deep craters with 3.5 mm outside diameter) for excitation as anodes in a d.c. arc. The arcing characteristics of the resulting mixtures were naturally determined largely by their high NaCl and BeO content. Consequently the above choice of an internal standard substance might, in spite of other favorable factors, be less desirable than one which participates in the arcing process more nearly like the minor constituents being determined, and more or less independent of the major constituents. However, it was realized that the success of the method would depend on whether, and to what extent, the vaporization and excitation of Mg, Ca, Sr, and Ba would parallel that of Be in the arc.

The details of the method just referred to will be presented elsewhere, while the present paper is intended to re-emphasize some fundamental features and sources of difficulty inherent in d.c. carbon arc methods of analysis. It is sufficient to mention at this point that the reproducibility of results, although usable, were not of the best. The sample was arced to complete consumption in order to integrate the radiation from all elements regardless of any fractional distillation in the arc. This proved to be particularly important, as in spite of all attempts to avoid it, sodium disappeared from the burning arc in about 50 seconds (at 6 amp.) leaving what appeared to be a pure Be-colored arc for another 30 seconds. Consequently the bulk of the internal standard substance was not participating in the arcing process at the same time as the Na and, as shown by later tests, the alkaline earths themselves. This explained the observation that intensity ratios from repeated spectra of the same sample sometimes differed by more than a normal amount even though the absolute intensity of the Be standard line remained fairly constant. Therefore, it appeared that the Be-standard produced a line of rather constant intensity when vaporizing from the arc as nearly pure BeO, but was not following the wider intensity fluctuations of the alkaline earth metals volatilizing earlier in the process along with NaCl.

In an attempt to secure more reproducible results than were possible with the Be standard, 0.25 percent Cr_2O_3 was introduced as a second intensity standard, into the same mixtures. It was found that in spite of its chemical difference, the excitation of Cr parallels very closely that of Mg, Ca, Sr, and Ba and therefore, is a much more reliable intensity standard for determining these elements. The use of Cr led to improved reproducibility in intensity ratios as compared to results based on Be in the same spectra.

It should be pointed out that an internal intensity standard can serve to calibrate intensities in two distinct ways:

(1) by following short period, or instantaneous variations, in the arc; and

* Paper presented at Optical Society of America Meeting, New York City, October 4, 1946.

(2) by fixing the general average intensity level of the integrated spectrum, as well as the sensitivity of the photographic recording.

The first of these *has a chance* to operate when intensity standard and analysis elements vaporize nearly parallel from the arc, as found for Cr and the alkaline earths. The second way is exemplified by the fractionated NaCl: BeO spectra mentioned above, in which a constant BeO content will follow variations in sample size and anything affecting the over-all density level of the photographed spectra—such as instrument adjustments, photographic material, and processing.

"CUT-OFF" PROCEDURE

Having observed that the different intensity standards used represented the two degrees of internal control mentioned in the preceding paragraph, it was decided to examine their behavior in more detail. For this purpose all standards were re-arced, and their spectra recorded only until a sudden drop in arc current was indicated. This corresponded to the visual disappearance of Na color from the arc column. This "cut-off" point was at about 50 seconds. Calibration curves were constructed from the spectra of these partially consumed standards in the usual manner described in previous publications.[1–3]

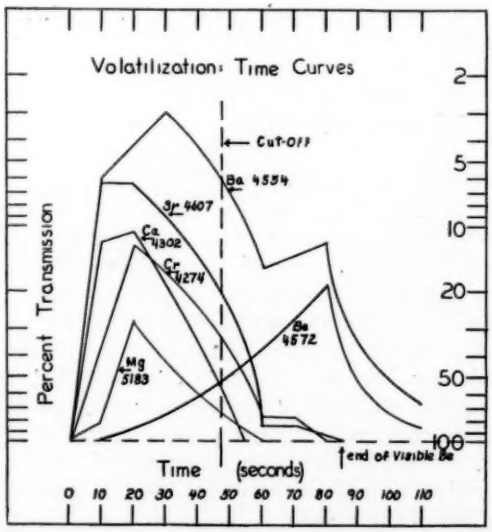

FIG. 1. Concentrations: Ca = 1.44 percent, Mg = 0.43 percent, Sr = 0.43, and Ba = 0.45.

Some general predictions regarding intensity ratios in these cut-off spectra, as compared with complete consumption spectra, can be made by reference to their intensity: time (fractionation) curves shown in Fig. 1 for one of the higher standards. This standard, as arced, contains 1.44 percent Ca, 0.43 percent each Mg and Sr, 0.45 percent Ba, 25 percent BeO, 0.25 percent Cr_2O_3, and balance NaCl—further combined with two parts carbon powder.

In the first place concentration calibration curves, based on Cr as intensity standard for the cut-off method will be displaced very little from the curve for complete consumption. This is because of the nearly parallel volatilization of Cr and the elements Mg, Ca, Sr, and Ba, whereby the greater part of each has volatilized during the pre-cut-off period ending at about 50 seconds. It is impractical to control the actual exposures so as to secure spectra of the arc only up to the point where line emission of these elements has fallen more nearly to zero.

It is clearly evident from the figure that the arcing process is interrupted at a time when both the analysis lines and the Be internal standard line intensities are undergoing their maximum rate of change *in opposite directions*. Intensity ratios of these lines are therefore very sensitive at this point and will be difficult to reproduce; exactly as found in practice. Furthermore, since the bulk of the analysis elements volatilize in the pre-cut-off period, while the bulk of the Be-intensity standard volatilized later in the arcing process, intensity ratios based on Be will naturally differ very considerable in the two procedures. For example in the cut-off method the intensity of Be will be much lower than in the complete consumption method.

COMPARISON OF "COMPLETE CONSUMPTION" AND CUT-OFF METHODS

Figures 2 through 5 are the actual concentration calibration curves constructed for Mg, Ca,

[1] L. W. Strock, "Quantitative spectrographic determination of minor elements in zinc sulfide ores," Am. Inst. Mining and Met. Eng. Tech. Publication No. 1866, 1945.
[2] L. W. Strock, *Spectrum Analysis with the Carbon Arc Cathode Layer* (Adam Hilger, Ltd., London, 1936), pp. 33–45.
[3] L. W. Strock, "Blank and background effect on photographed spectral lines," J. Opt. Soc. Am. **32**, 103–116 (1942).

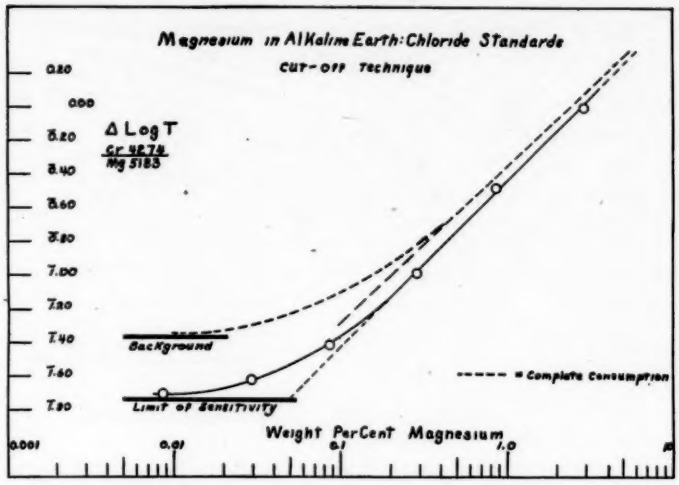

FIG. 2.

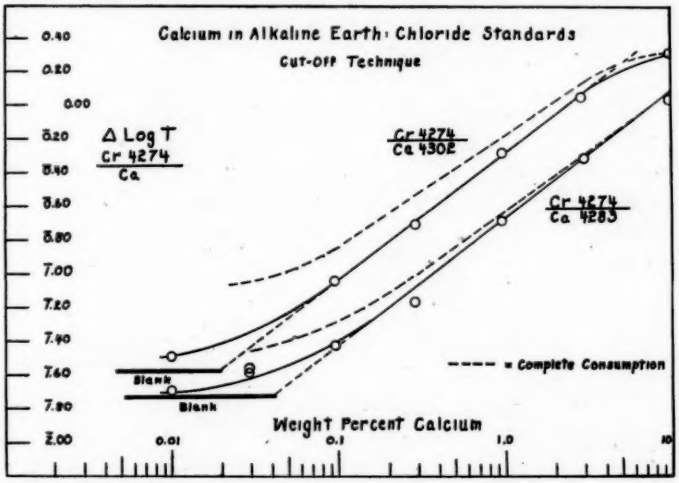

FIG. 3.

Sr, and Ba using Cr as internal standard for the cut-off technique; that is, whereby all light obtained in the pre-cut-off period is integrated into one spectrum. For comparison the curves obtained by complete consumption are drawn on the figures as dashed lines. Figure 2 illustrates the curves obtained for the line Mg-5183.618 using Cr-4274.803 as internal standard. The two curves are close together at higher concentration. Two other features of this curve pair are important: (1) the much lower background for the cut-off curve, as a result of extinguishing the arc while burning at its initial lower temperature (the backgrounds differ by more than a factor two, with a resulting extension of the method to lower concentration for the cut-off method), and (2) the slight difference in slope of the linear portions of the two curves—the cut-off one having a greater slope. Figure 3 illustrates the corresponding curves for the lines Ca-4302.527 and Ca-4283.010 using Cr-4274.803 as intensity standard, while Fig. 4 is derived from the two strontium lines Sr-4607.331 and Sr-4811.881. Figure 5 is derived from the two barium lines Ba-4554.042 and Ba-4934.086. All the above curves, based on a Cr line as intensity standard, exhibit the same

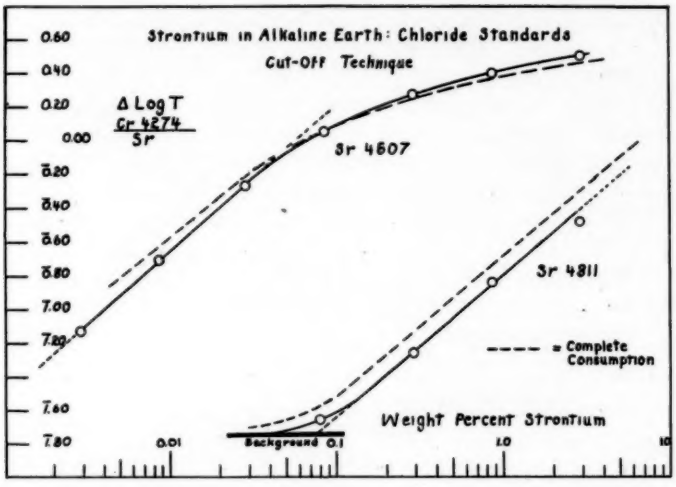

FIG. 4.

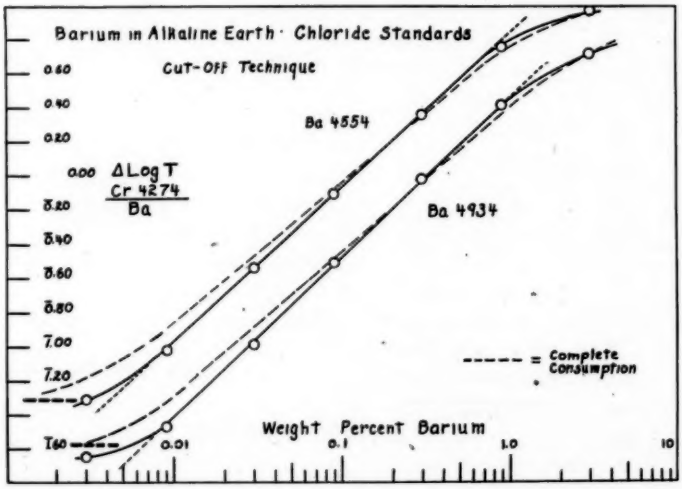

FIG. 5.

features for the cut-off method as compared with the complete consumption one mentioned above under Mg. This is most clearly shown, however, by Ba where the curves for both methods intersect.

Figures 6 through 9 are the corresponding curves for the same lines (in the same spectra) using Be-4572.671 as intensity standard. Figure 6 is derived from the line Mg-5183.618 using Be-4572.671 as intensity standard. The marked displacement of the cut-off curve from the complete consumption one in this case, is a result of the low Be-line intensity in the cut-off method spectra wherein Be was very incompletely consumed. Accordingly, the intensity of the Be line relative to the Mg line (as measured by $\Delta \log T$) is much less than in complete consumption spectra. Further, the slope of the cut-off curve *is less* than the complete consumption curve when Be is used as intensity standard, while the slope of corresponding curves based on chromium is greater. This same feature is shown by the curves for the three other elements studied.

Figure 7 is derived from the two calcium lines Ca-4302.527 and Ca-4283.010 using Be-4572.671 as intensity standard. Here, as with Mg the cut-

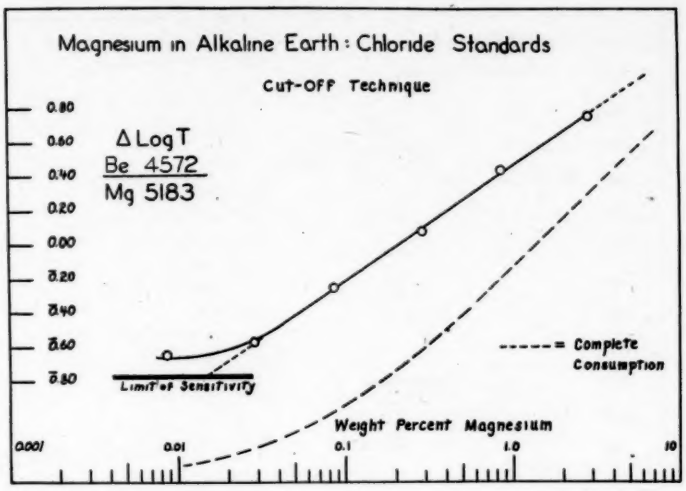

FIG. 6.

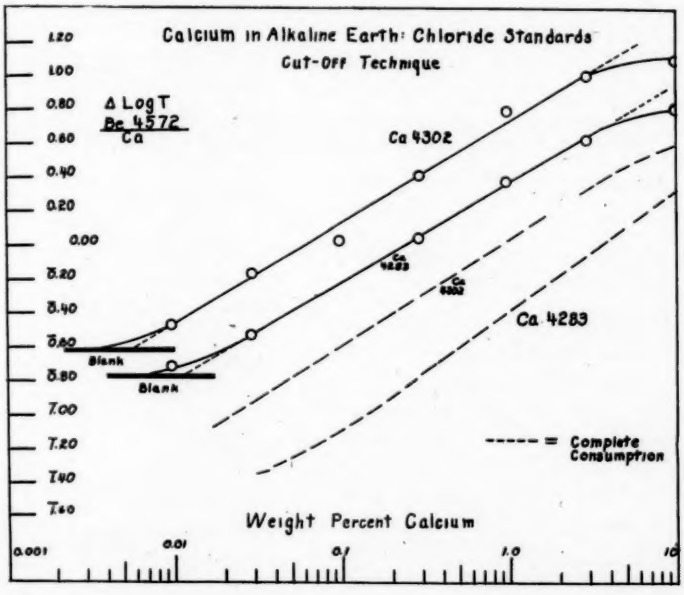

FIG. 7.

off and complete consumption curves are separated by a Δ log T value determined largely by a weakening of the Be-line intensity in spectra of the incompletely consumed standards in the cut-off method. Again the cut-off curve slope is lower than found for the completely consumed sample.

Another important feature shown by the Be : Ca curves is the flatter shoulder on the cut-off curves. The explanation of this will be given below, and in the meantime it is illustrated more effectively by the Sr and Ba curves.

Figure 8 illustrates the results obtained for the two strontium lines Sr-4607.331 and Sr-4811.881 using Be-4572.671 as intensity standard. It should be noted that these two Sr lines supplement each other, and between them permit the use of the method over a concentration range 0.001 to 5.0 percent Sr. Otherwise, the greater

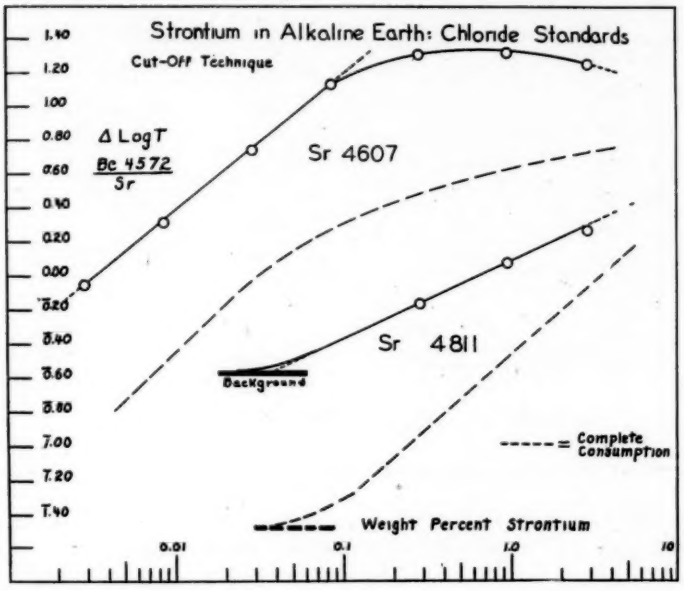

FIG. 8.

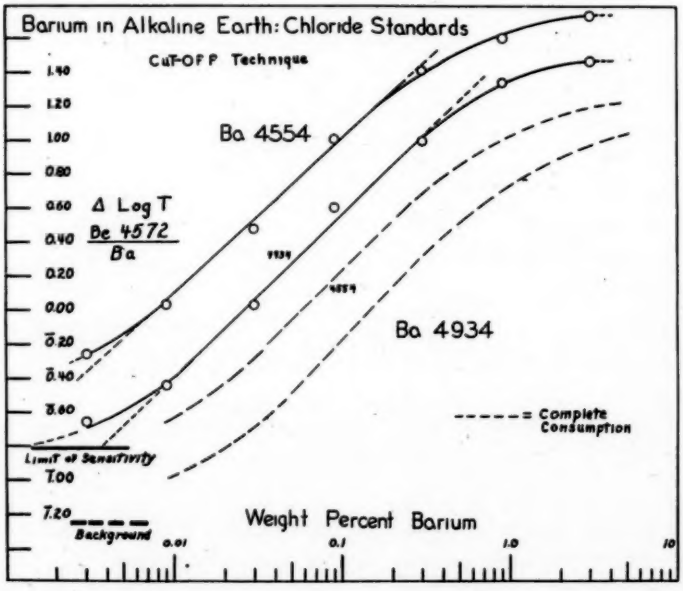

FIG. 9.

abruptness with which Sr line intensity ceases to increase with concentration in the cut-out method is clearly evident, as well as the lower rate of Sr intensity increase with concentration below the shoulder.

Figure 9 shows the curves derived from the two barium lines Ba-4554.042 and Ba-4934.086 using Be-4572.671 as intensity standard. Here the slopes of the middle portions of the curves are very similar, although as with Mg, Ca, and Sr,

the cut-off curve slope is actually less. The more rapid flattening of the cut-off curve at higher concentrations is evident.

The differences and similarities of the curves for the two procedures pointed out above are:

(1) *The small displacement* of the curves *when using* Cr *as intensity standard* (except for the resulting difference in background) is primarily caused by a slightly *greater slope* of the cut-off method curves.

(2) *The large displacement* of the curves *when using* Be *as intensity standard*, is primarily caused by a marked decrease in Be line intensity resulting from interrupting the arc before the internal standard itself had volatilized. In this case, the cut-off method curves have *lower slopes*.

EFFECT OF FRACTIONAL DISTILLATION ON CONCENTRATION CURVES

All of the features pointed out for the above curves are attributable to the fractional distillation of the various constituents from the arc in the order already illustrated by the volatilization curves of Fig. 1. A comparison of the Cr and, for example, Ca volatilization curves shows that these elements distill from the arc very nearly in a parallel manner—neither have completely volatilized at the cut-off point when Na disappears from the arc. In the 2.87 percent Ca standard illustrated, Ca has volatilized to a greater degree than has Cr. *The greater slope of the cut-off curve* therefore indicates; that *at lower Ca concentrations, an increase in the degree of volatilization of Cr* over the degree achieved at higher concentrations has occurred. In other words, as the alkaline earths decrease, Cr vaporizes earlier in the arcing process.

On the other hand the *lesser slope* of the cut-off method curve for the Be-standard indicates; that *at low alkaline earth concentrations, Be volatilization has been retarded*, and takes place largely after the cut-off point when Na has disappeared. This obviously leads to lower Be-line intensities for the low standard (and correspondingly higher intensity ratio values) and a flattening of the concentration curves in the manner actually found.

That this explanation is correct, is illustrated by the volatilization curves for a low standard shown in Fig. 10. This standard, as arced, contains 0.048 percent Ca, 0.015 percent each of Mg, Sr, and Ba, 25 percent BeO, 0.25 percent Cr, and

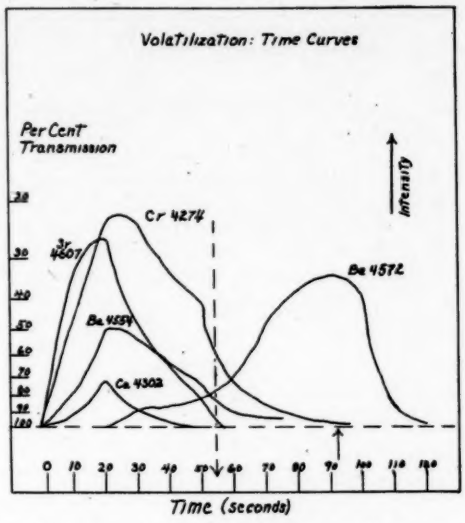

FIG. 10. Concentrations: Ca = 0.096, Mg = 0.029, Sr = 0.029, Ba = 0.030.

the balance NaCl—further combined with 2 parts carbon. The Cr volatilization has moved farther into the pre-cut-off field, while Be volatilization has moved farther into the post-cut-off field.

SELF-ABSORPTION DIFFERENCES IN THE TWO METHODS

One other feature of the calibration curves which varies with the two procedures is the observed more rapid flattening of the cut-off curves at higher concentrations. This can be explained in terms of self-absorption, of the alkaline earth lines themselves, which effects the concentration curve in the manner illustrated by Fig. 11. Here the vapor density of the alkaline earth metal in the arc column is sufficiently high in the pre-cut-off period to cause marked self-absorption of their lines. Since, at higher concentrations, considerable alkaline earths remain in the arc after the cut-off point, and continue to distill off with the Be-residue, the intensity of alkaline earth lines, now less effected by self-absorption, are sufficient to build up the line densities obtained earlier in the arcing process when self-absorption had occurred. Accordingly, the complete consumption curve is the sum of the cut-off one and the cut-off residue. Note that this is an actual

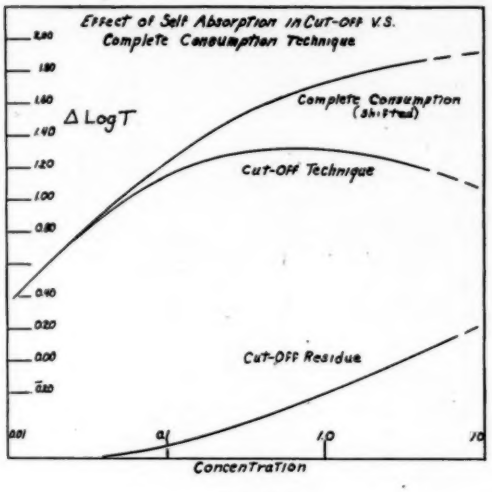

FIG. 11.

experimental curve constructed for Sr-4607.331, the complete consumption one merely being shifted to coincide with the cut-off curve at a point below the self-absorption effected concentration level. Also note that no Sr remains in the cut-off residue under 0.05 percent Sr content.

CONCLUSION

The data presented in this paper clearly contradict our initial hope that the cut-off technique would represent an improvement over the complete consumption technique first developed, and permit the use of Be as an internal standard, even though very incompletely volatilized. It has been shown that sufficient Be does participate in the arcing process during the Na volatilization period to provide an intensity standard line, but that the cut-off point, determined by externally recognizable features of the arc, lies in a range where intensity ratios are changing so rapidly with time as to be poorly reproducible. The results of our study undertaken to examine this possibility, do illustrate many basic features and sources of difficulty in d.c. carbon arc methods of spectrochemical analysis; and for that reason have been presented here in some detail.

Thermal and X-Ray Analyses of Some Common Phosphors

RUDOLPH NAGY AND CHUNG KWAI LUI
Research Department Lamp Division, Westinghouse Electric Corporation, Bloomfield, New Jersey
(Received October 10, 1946)

Thermal analysis studies of some common phosphors were made. The exothermic and endothermic reactions of the phosphors were correlated with the fluorescent output and x-ray diffraction patterns. The maximum output of fluorescence was obtained from well-crystallized phosphors as manifested by sharp x-ray diffraction lines. Thermal analyses in conjunction with x-ray analyses have been used to elucidate the changes which occur in the formation of phosphors and in determining their proper firing temperatures.

INTRODUCTION

IN view of the fact that thermal analyses have been found to be of considerable value in studying changes in mineral structures,[1] it would seem very desirable to apply this method in the analysis of the various steps in the formation of a phosphor.

The method of thermal analysis as perfected by Norton[2] has been used to study the dehydration, formation, inversion, crystallization, and melting of compounds. It would appear, therefore, that this method may be another convenient way by which the mechanism of fluorescence can be investigated. The various exothermic and endothermic reactions in the phosphor can be studied and the corresponding crystal configuration determined by means of x-ray diffraction. Thermal analysis studies of fluorescent compounds in conjunction with x-ray diffraction analyses are shown here to provide a means of determining the proper formation temperatures of phosphors.

APPARATUS

The schematic diagram of the apparatus used for thermal analyses shown in Fig. 1 was similar to that used by Norton.[2] A program controller was used to increase the temperature of the furnace at a constant rate. The actual temperature of the furnace was determined by placing a sample of alumina in the furnace and imbedding platinum-10 percent rhodium thermocouple *a* in the mass. A portable potentiometer *A* was used to record the temperature. A differential thermocouple *b* was imbedded in the alumina and the raw ingredients of the phosphor so that only the difference in temperature of the two samples would be measured. A sensitive potentiometer *B* capable of reading to a hundredth of a millivolt was used for these measurements.

A Picker x-ray diffraction unit with an iron tube operated at 30 kv and 10 ma was used for the x-ray analyses. A small Debye-Sherrer powder camera having a radius of 3.59 cm was employed for identification purposes. A back reflection camera was used to study the shift in lattice spacings of the substances.

METHOD

The exothermic and endothermic reactions occurring in the formation of the various phosphors were studied by slowly heating the unfired materials in a thimble size platinum crucible. These were placed in an electric furnace and the temperature changes in the samples were ob-

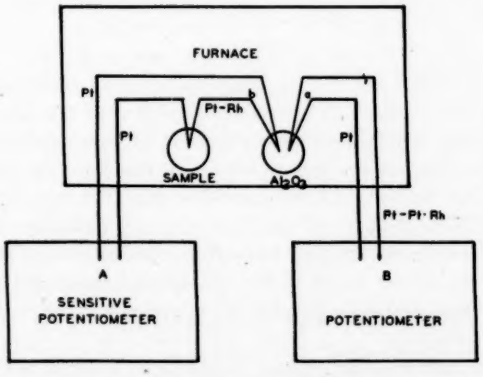

FIG. 1. Schematic diagram of apparatus for thermal analysis of phosphors.

[1] S. Speil *et al.*, Bur. Mines Tech. Pap. 664 (1945).
[2] F. H. Norton, J. Am. Ceram. Soc. 22, 54–64 (1939).

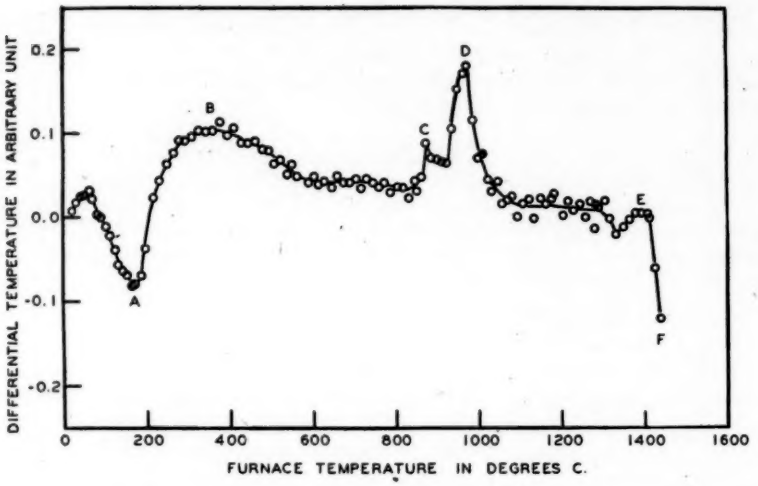

FIG. 2. Thermal analysis curve for zinc silicate.

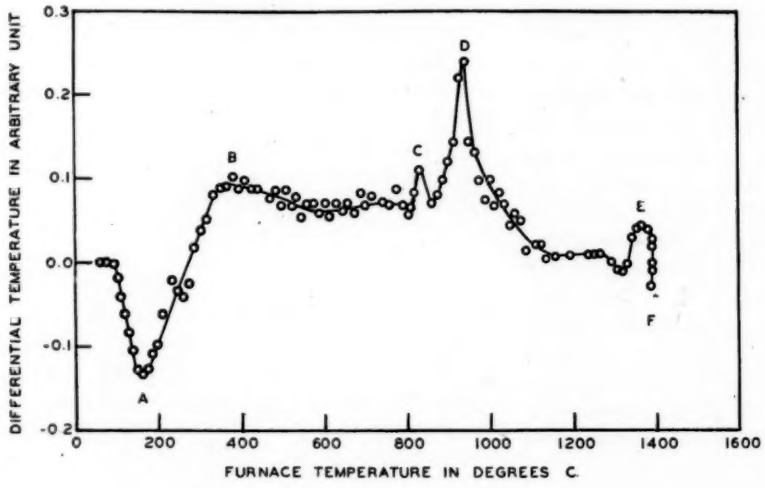

FIG. 3. Thermal analysis curve for zinc beryllium silicate.

served. The rate of increase in temperature of the furnace was thirteen degrees per minute. Lower and higher rates of heating were also tried but were not found as satisfactory. The temperature of the furnace was taken every five minutes and the differential temperature was taken every minute. Thermal analyses curves indicating the exothermic and endothermic reactions were obtained by plotting the differential temperature expressed in millivolts against the furnace temperature.

Samples for x-ray diffraction studies were taken out of the furnace just before and after any significant exothermic or endothermic change occurred in the sample and the crystal structure of the substance determined. When the shift in lattice spacing was to be measured by means of a back reflection camera the sample to film distance was set at 5.10 cm. Four pictures each requiring an eight-hour exposure were taken on the same film.

RESULTS AND DISCUSSION

1. Zinc Silicate

The raw materials for zinc silicate were prepared by ball milling 1.2 mole zinc oxide, 1 mole silicic acid, and 0.024 mole manganous carbonate.

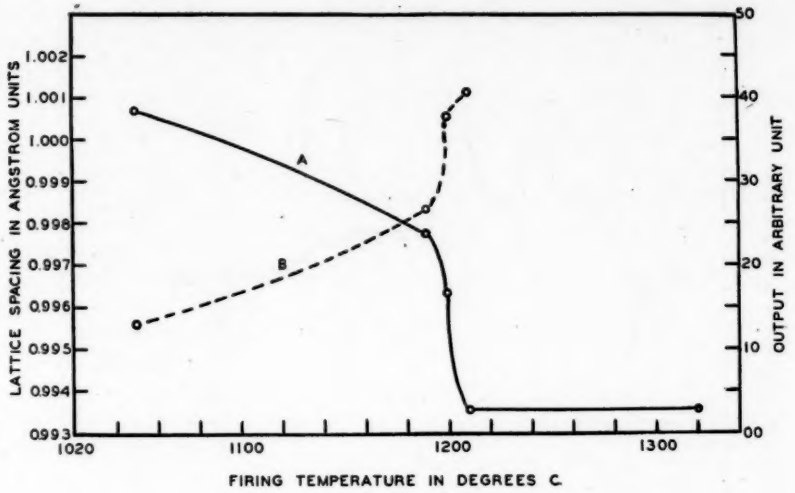

FIG. 4. Effect of firing temperature of zinc beryllium silicate on: (*A*) lattice spacing, (*B*) output.

A typical thermal analysis curve is shown in Fig. 2. The first endothermic reaction at point *A* indicates a loss of water from silicic acid. Heat would be required to remove both the absorbed and chemically bound water. Point *B* has not been satisfactorily explained but may be caused by a more rapid diffusion of heat into zinc oxide and silica than into the alumina. The preliminary heating has a tendency to make the mass very porous. X-ray diffraction pictures taken before and after point *C* showed no change in crystal structure. The x-ray patterns were mainly those of zinc oxide. The silica evidently was in an amorphous state and gave no contribution to the x-ray pattern. The exothermic point *C* may represent a change in silica from one form into another. Point *D* corresponds to the formation of zinc silicate. X-ray diffraction patterns for samples taken between *D* and *E* correspond to ortho silicate structure. Point *F* is the melting temperature of the zinc silicate. Thus a zinc silicate phosphor can be made by heating the ingredients anywhere between 1000°C and 1400°C with favorable results. The melting point *F* and formation point *D* can be slightly lowered by varying the amount of manganese which acts as a flux and also by the particle size of the raw ingredients.

2. Zinc Beryllium Silicate

The materials for zinc beryllium silicate were prepared by ball milling together 7.4 moles zinc oxide, 1 mole beryllia, 5 moles silicic acid, and 0.5 mole manganous carbonate. The thermal analysis curve Fig. 3 is very similar to that of zinc silicate. X-ray diffraction patterns of the samples taken between *D* and *E* have a zinc silicate structure. However, the back reflection x-ray diffraction lines show a contraction of lattice spacing at high temperatures as shown in curve *A* of Fig. 4. This contraction of lattice spacing at high temperatures can be explained by the fact that the smaller beryllium ions substitute for the larger zinc ions. The sample fired at 1050°C fluoresced green and the color changed gradually to yellow, orange, and pink as the firing temperature increased. It would appear, therefore, that in the heating of the raw materials, zinc silicate was formed first and at the higher firing temperatures, the beryllium slowly entered into solid solution with the zinc silicate forming zinc beryllium silicate. The output of the product, as shown by curve *B* of Fig. 4, was closely related with the contraction of the lattice spacings. Both the output and lattice spacings change very rapidly in the short temperature range between 1190°C and 1210°C and, therefore, the optimum firing temperature would be very critical in this region.

3. Magnesium Tungstate

The raw materials for magnesium tungstate were prepared by ball milling together 1.1 moles magnesium oxide and 1 mole tungstic acid. The

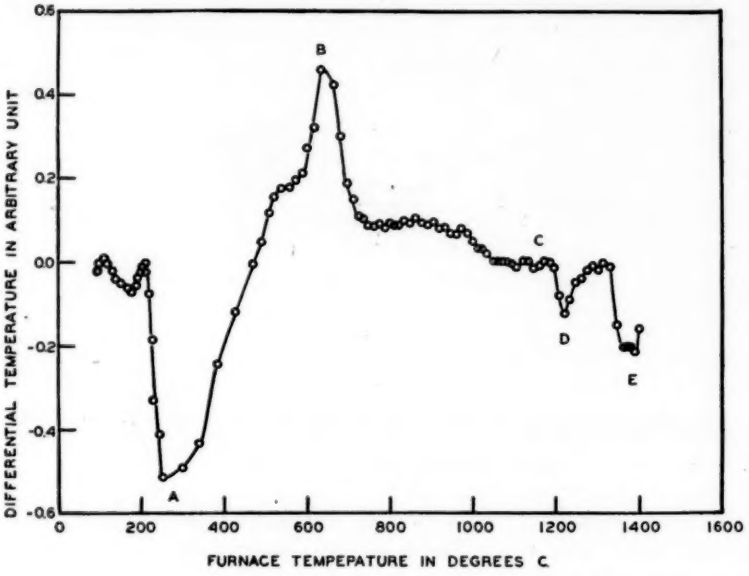

FIG. 5. Thermal analysis curve for magnesium tungstate.

thermal analysis curve is shown in Fig. 5. Point *A* represents an endothermic process which in this instance would be dehydration of tungstic acid. The reaction between the tungstic oxide and magnesium oxide occurs at point *B*. The fluorescent output increased as the heating proceeded so that at point *C* the maximum fluorescence was obtained. X-ray diffraction patterns for samples between points *B* and *C* indicate a magnesium tungstate with the monoclinic structure. The diffraction lines were very sharp indicating that the substance was well crystallized. Between the temperatures of 1200°C and 1250°C a complete loss of fluorescence was observed and the x-ray diffraction patterns revealed an entirely different crystalline structure. Fonda[3] assumed that this substance was Mg_2WO_5 because he obtained this new structure with 1.93 mole magnesium oxide or higher to 1 mole tungstic acid and not when lower ratios of magnesium oxide were used. X-ray diffraction patterns of phosphors with 1.1 mole magnesium oxide to 1 mole tungstic acid fired at 1280°C matched exactly with the proposed Mg_2WO_5 pattern. If Fonda's assumption were correct an excess of one mole of tungstic oxide would have been produced according to the following formula:

$$2MgWO_4 = Mg_2WO_5 + WO_3.$$

However, the diffraction patterns do not show any WO_3 and furthermore, no change in weight has been detected by heating the fluorescent materials to 1280°C thus showing that WO_3 has not been lost.

The non-fluorescent magnesium tungstate may be a high temperature form of the compound. Compounds with ratios of the ionic radii of cation and anion very near to the critical ratio of 0.22, 0.41, or 0.73 may exist with two different coordination numbers. A very common example of this phenomenon can be seen in the two forms of $CaCO_3$, calcite and aragonite, in which there is a coordination of six and eight, respectively. Since the ratio of cation to anion radii for magnesium tungstate is 0.46 it is reasonable to believe that there exist two crystal forms having different coordination numbers.

In the higher temperature forms of compounds the ionic distance between the cation and anion would be increased, therefore, the total energy of the crystal should be increased. The endothermic reaction of $MgWO_4$ begins at 1200°C and clearly indicates an absorption of energy which would substantiate the theory of a higher

[3] G. R. Fonda, J. Phys. Chem. **48**, 303 (1944).

temperature form for this compound. Point *E* which represents the melting point of the phosphor has been found to be dependent upon the ratio of the magnesium oxide to tungstic acid.

4. Cadmium Borate

The thermal analysis of a mixture of 2 moles of cadmium oxide, 1 mole of boric oxide, and 0.032 mole of manganous carbonate gave the thermal curve shown in Fig. 6. Point *A* represents the dehydration. Point *B* indicates a violent reaction to form cadmium borate. Fluorescence does not occur until point *C*. The x-ray diffraction patterns show no change with increase of heating until melting point *D*. Slight variations in the temperature at which these points occur were dependent upon the composition of the phosphor.

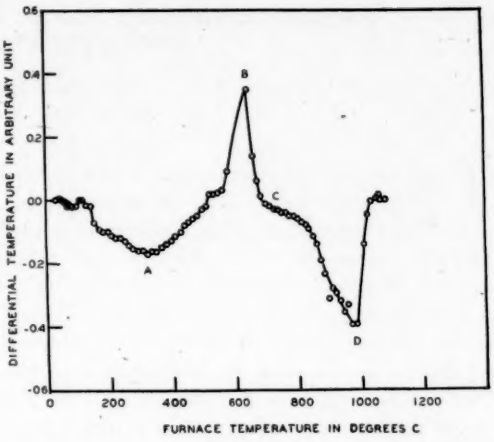

FIG. 6. Thermal analysis curve for cadmium borate.

CONCLUSION

Attempts have been made to correlate the various exothermic and endothermic reactions of phosphors observed during heating, with the changes in crystal structure. These results indicate that the optimum heating temperatures for maximum fluorescence could be determined from the thermal and x-ray analyses of the phosphor. All phosphors have definite and well formed crystal structures as evidenced by their sharp x-ray diffraction lines.

The authors wish to express their appreciation to Dr. Joseph A. Pask for his helpful suggestions on thermal analysis, and to other members of the Research Department of the Westinghouse Electric Corporation at Bloomfield, New Jersey.

An Autocollimator for Precise Measurement of the Flange Focal Distance of Photographic Lenses*

M. G. Townsley and P. C. Foote
Bell & Howell Company, Chicago, Illinois

THE flange focal distance of a lens is defined as the distance from the mounting surface to the back focal plane. This distance is of considerable importance in photographic lenses where it is necessary to have interchangeability of lenses from one piece of apparatus to another, or of a multiplicity of lenses on a single piece of apparatus. It is usual to measure the flange focal distance with a microscope focused first on the flange plane and then on the image of an infinitely distant object. Either an object at a physically great distance or a reticle in the focal plane of a collimator may be used. It is also possible, although not so commonly done, to measure the flange focal distance by means of a reticle in the back focal plane and a collimated telescope.

It may be pointed out that the flange focal distance of a lens may not necessarily be identical with the flange to film distance of the camera with which it is intended to be used. In the first place, the "back focal plane" may not be a plane, but a curved focal surface. The flange focal distance, at least in the instrument to be described here, is measured on the axis, and it may be necessary to introduce a correction to take into account compromise focusing due to curvature of field. Also, it is usually desirable, in the case of short focus lenses having fixed mounts, to set or face the lenses for the hyperfocal distance.

Some time ago, it became necessary for one of the authors to measure the flange to film distance and the flange focal distance of a group of lenses under conditions such that the back focal plane was inaccessible. Specifically, it was desired to determine the position of film running in a motion picture camera. The instrument which was developed for this purpose has been refined and improved to broaden its usefulness, and is the subject of the present paper.

Because of its history, the instrument has come to be known in our laboratories as the "Film Focus Collimator." The name "microfocuser" has been suggested as more appropriate, to denote that it is a tool which is primarily used for accurately focusing lenses. Basically, the instrument makes it possible to use a standard micrometer spindle to measure the position of an optical image by providing a means for accurately determining when the spindle face of the micrometer is in the plane of the image.

The optical system is shown schematically in Fig. 1. The optical and mechanical arrangement is designed to be flexible and precise. As shown in Fig. 1, the microfocuser has a lamp, a source reticle, an axially adjustable objective, and a micrometer. The lens under test is located between the objective and the micrometer. Between the source reticle and the objective is interposed a beam-splitter which partially reflects any light entering the objective into an eyepiece tube at right angles to the main body tube. A reticle in the eyepiece tube defines the focal plane of the eyepiece, and is placed with respect to the beam splitter so that it is imaged in the plane of the source reticle. Thus, any imaging ray entering the objective in such a manner as to form an image in the plane of the source reticle will also be imaged in the plane of the eyepiece reticle. The location of the eyepiece reticle is described in this way rather than by simply saying that it is in the principal focus of the objective because, for some applications of the instrument, the objective is displaced axially from its infinity focus position.

Figure 2 shows the form in which the instrument has been used in our laboratories, assembly, and inspection departments. The lamp is a 1000-watt motion picture projection lamp, placed in the lamphouse at the extreme left of the figure and cooled by a suitable fan. The source reticle is placed at the left end of the body tube and consists of crossed pairs of closely spaced lines ruled in opaque silver. A cubical glass block

* Presented at the thirty-first annual meeting of the Optical Society of America, October 3–5, 1946, New York, New York.

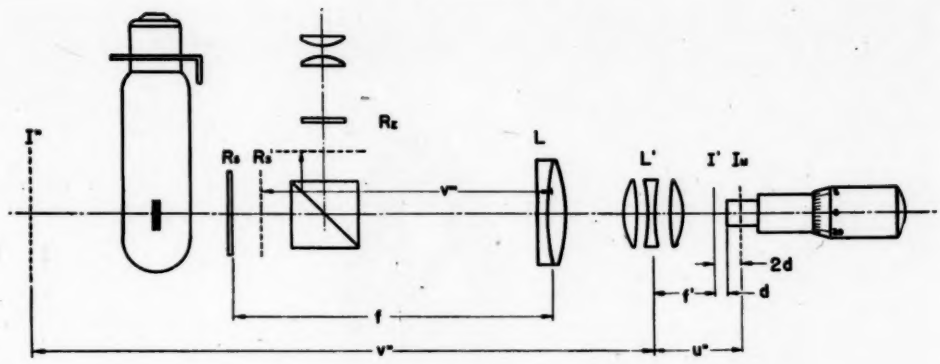

FIG. 1.

with a diagonal interface coated with a semi-reflecting coating forms the beam-splitter. The objective is a doublet of 18 inches focal length having a clear diameter of 2½ inches. The eyepiece tube is folded for convenience, and is equipped with a dark line eyepiece reticle and a one-inch eyepiece. At the right of the body tube is seen the mount for the auxiliary lenses which will be described later, and the mount for the lens under test and the micrometer spindle. The entire unit is mounted on an optical bench consisting of a pair of heavy cylindrical bars supported on a pipe and cast iron framework, this form of optical bench being standard in our organization.

The diameter of the objective is large enough to clear all lenses normally supplied, and sufficient illumination is provided for the smallest aperture to be used. The focal length of the collimator objective is longer than any lens normally tested, to obtain adequate magnification.

The mount for the lens to be tested has appropriate means for attaching the lens to be tested, and for locating the lens from the flange or shoulder from which the flange focal distance is to be measured. A micrometer spindle is attached to this mount, coaxial with the lens to be tested and adjusted so that the distances indicated on the micrometer scale are measured from the seat against which the mounting flange rests. The face of the spindle is polished accurately square with the spindle axis, and flat within one or two wave-lengths.

To measure the flange focal distance, the lens to be measured is placed on the lens mount, facing the objective L (Fig. 1). The objective L is set to the position which produces an image of the source reticle R_s at infinity, so that the light between the objective and the lens being measured is collimated. The lens L' under test, will then produce an image of the source reticle in its back focal plane, at I'. The face of the micrometer spindle acting as a mirror, and being at a distance d from the focal plane I', will form an image I_m' at a distance $2d$ from the focal plane. This reflected image now becomes an object at a distance $u''=f'+2d$ for the lens L' being measured, which produces an image I'' at a distance v'' in accordance with the classical equation for object and image. This image I'' in turn, is an object for the objective L which forms an image R_s' at a distance v'''. It is obvious that, if the micrometer spindle is moved to reduce d, u'' approaches f', v'' approaches infinity, and v''' approaches f, bringing the resultant image R_s' of the source reticle closer to the eyepiece reticle R_e. It is also clear that $2d$ approaches zero twice as fast as the micrometer spindle face approaches the focal plane of L'. It is this doubling of the focusing effect which gives the instrument its extreme sensitivity. When the source reticle image comes into focus in the eyepiece reticle plane, the flange focal distance may be read directly from the micrometer spindle. Parallax elimination between the source reticle image R_s' and the eyepiece reticle R_e is helpful, but not essential, for accurate focusing.

The relationship between the movement of the micrometer spindle and the movement of the eyepiece image may be more elegantly developed

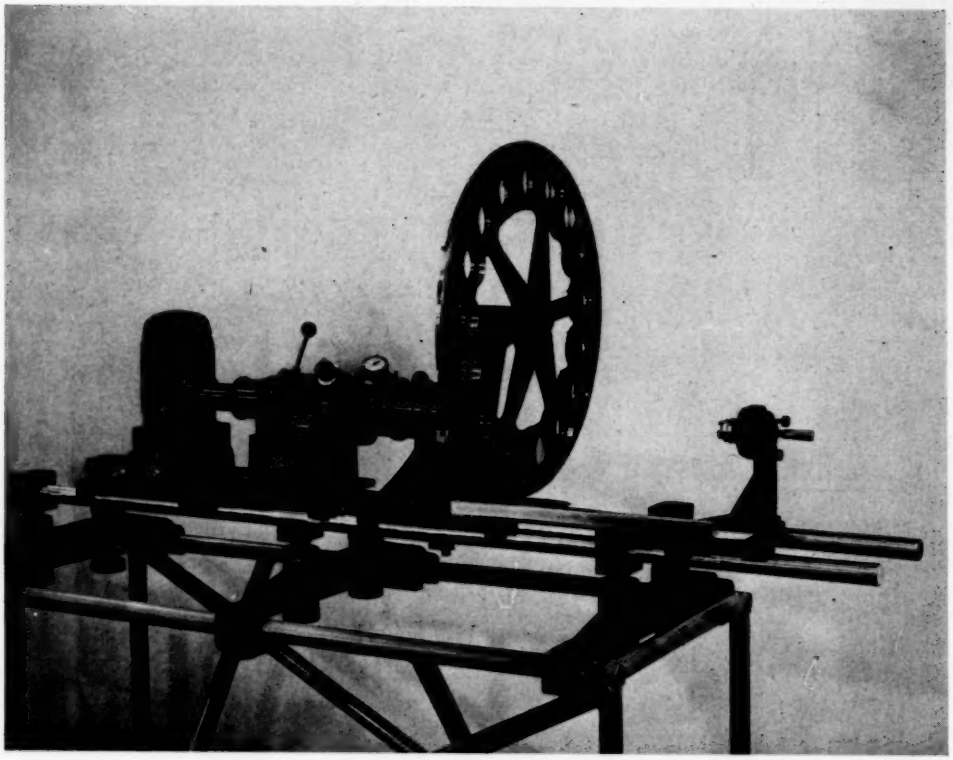

FIG. 2.

by considering the lenses L and L' as components of a single lens having a focal length F. Then

$$1/u''+1/v'''=1/F,$$
$$-u''^{-2}du''-v'''^{-2}dv'''=0,$$
$$dv'''/du''=v'''^2/u''^2;$$

we may call dv'''/du'' the longitudinal magnification but, $M=v'''/u''$ where M is the lateral magnification whence, $dv'''/du''=M^2$. As we have already seen, if d is the movement of the micrometer spindle,

$$du''=2dd,$$

and therefore,

$$dv'''/dd=2M^2,$$

so that the microfocuser has twice the longitudinal magnification of a telescope or microscope of equal lateral magnification.

The instrument as described and as shown in Fig. 2 has been used to measure the flange focal distance of a considerable number of motion picture camera lenses for the purpose of facing the flanges to a fixed standard flange focal distance and for inspecting the lenses after facing.

In addition to facing lenses to fixed flange focal distances, the instrument is equipped for the calibration of focusing scales. Two methods are used to effect this calibration: either the objective is racked in to produce an image at the desired distance, or the objective is set at infinity and an auxiliary lens is used.

In the first method, the objective is racked in toward the source reticle by a predetermined amount, to produce a virtual image at a distance v to the left of the objective, which then becomes an object for the lens under test at a distance $D=v-u+A$, where A is the distance between the reference plane of the lens being tested and the source reticle. In motion picture practice,

D is always measured from the film plane, so that A is the distance from the source reticle to the nominal position of the micrometer spindle. The micrometer spindle is set to the flange focal distance with the lens focused in its infinity position, and the lens is moved by means of its focusing mount until the source reticle is sharply imaged in the eyepiece reticle plane. The lens under test is now focused for the distance D and the point may be suitably marked on the focusing scale.

To facilitate accurate repetition of the settings of the objective in calibrating focusing scales, it has attached to the lens cell a precise dial indicator reading in .001". The nose of the indicator bears, through a hinged block, on the end of a rod which is secured to the main body tube, parallel to the axis of the tube. A series of these rods is mounted in a pair of rings which rotate around the tube so that any one of the rings may be brought opposite the indicator. The lengths of the rods are adjusted to give a desired series of positions of the objective and corresponding image distances D. If the indicator is adjusted to read zero when in contact with the infinity rod with the objective accurately focused for infinity, any desired distance may be set off by rotating the nest of rods to bring the desired rod opposite the indicator and adjusting the objective tube to again bring the indicator to zero. All of the commonly engraved distances from infinity to 15 feet are provided for in this way.

For distances closer than 15 feet, the movement of the objective becomes prohibitively large, and it is necessary to resort to the use of auxiliary lenses. The objective is set to the infinity position and auxiliary lenses are introduced between it and the lens being calibrated. A negative element causes a virtual image of the source reticle to be presented to the lens under test as though there were a real object at the required distance. A separate auxiliary lens is required for each distance D. It would be possible to use fewer auxiliaries and readjust the position of the objective for some distances, but this does not lend itself well to production use.

Figure 3 shows the optical layout of the system with an auxiliary lens in place for calibrating the focusing scale at a distance D. The distance

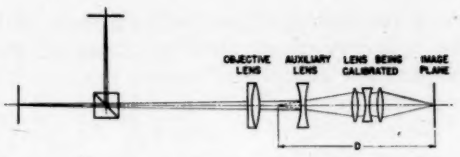

FIG. 3.

is measured from the image plane, in the same way as when the objective is racked in. It is necessary, therefore, that the position of the image plane be fixed and that the auxiliaries be located to produce their virtual images at the proper distance from the image plane. A distance from the auxiliaries to the image plane has been arbitrarily chosen, large enough to give adequate clearance for the longest lens to be inserted for calibration. For convenience, the auxiliaries are mounted on a large wheel, carrying enough stations to supply one for each distance to be calibrated from 15 feet down to one foot. The auxiliary lenses are more than one foot from the image plane, so a positive lens is used for this shortest distance, forming a real image of the source reticle in the correct position.

Accurate readings require considerable accuracy in the construction of the component parts of the instrument. To avoid astigmatism, it is necessary that the objective and the beam-splitting prism block be accurately made and free from strain or warpage due to cementing or mounting. The unused faces of the block are ground and blackened to eliminate stray reflections. For maximum sensitivity, the collimator reticle consists of crossed pairs of lines ruled in opaque silver and separated by a distance which will make them subtend approximately 3 minutes in the eyepiece. The eyepiece reticle is preferably a dark 60° cross.

Original collimation is accomplished with the aid of an accurately flat autocollimating mirror and an auxiliary telescope, by first using the telescope to bring the eyepiece and source reticles into common focus, and then using the mirror to establish the infinity setting of the objective. The accuracy of the auxiliary lenses and their setting is critical. Each lens is mounted in an adjustable sleeve so that it may be shifted to make the final focusing adjustment, but most of the adjustment was made by polishing the

lenses stronger or weaker to bring them within the relatively short focusing range of these mounts. The difficulty of making this correction increases directly with the focal length.

The adjusting of the power of the auxiliary lenses and their final setting in the mounts was accomplished by using a long focus, high aperture, lens in the position of the lens to be tested, and a high power microscope on the image. A suitable target was placed at the required distance, and the microscope adjusted to focus on its image. The microfocuser and the auxiliary were then switched in, and the auxiliary adjusted to bring the image of the source reticle into the proper plane.

The accuracy with which settings of the instrument can be made depends on the aperture and focal length of the lens being tested. Within the range ordinarily used for motion pictures, it is usual to repeat results between observers within $\pm.0005$ inch per inch of equivalent focal length. If large aberrations exist in the lens being tested, it is obviously impossible to make accurate readings because there is no sharp image on which to focus. The magnification of the system is sufficient to show up poorly corrected or defective lenses.

It was said in the introduction that the predecessor of the present instrument was originally developed during an investigation into the location of film in a running motion picture camera. The present instrument is still used occasionally for the same purpose. It is relatively easy to make accurate settings using the surface of the film itself as a "reflecting surface," instead of the micrometer spindle. In this case, the film does not form an actual reflected image, but becomes a receiver of an out of focus image and an out of focus diffuse radiator of this image, so that the doubling of the defocusing effect is present in the same way. Measurement of the film position is made by a transfer method. The instrument is used to observe the image on the film with the camera lens set to infinity, and the focusing is done by adjustment of the objective of the microfocuser. The camera lens is then transferred to the micrometer adapter, and, without disturbing the setting of the microfocuser, the micrometer is advanced to bring the image into focus in the eyepiece reticle plane. The micrometer setting is then the distance from the mounting flange to the film plane.

The instrument which has been described here has proved in practical use to be more accurate and easier to use than conventional microscope methods. It has been found possible to train unskilled operators to use the instrument within a few minutes to the point where they obtained better repeatability of settings than could be obtained by the most skilled operator using a conventional microscope. In one test, a group of four unskilled operators showed less scattering of settings among the group than a single skilled operator using a microscope showed among his own settings.

Telescope Objective without Spherical Aberration for Large Apertures, Consisting of Four Crown Glass Lenses

GORDON S. FULCHER
4712 Alton Place N.W., Washington, D. C.

(Received August 26, 1946)

A telescope objective of 10-cm focal length may be produced with practically zero spherical aberration for all wave-lengths of the visible spectrum and for apertures up to $f:2$ by combining four thin lenses of crown glass ($n=1.523$) of the same converging power. The lens surface curvatures approximate those of minimum aberration lenses but are in integral half diopters as follows: 1st lens, 2.5 and 0 diop.; 2nd, 4 and -1.5 diop.; 3rd, 6 and -3.5 diop.; and 4th, 7.5 and -5.0 diop. Detailed results are given for the linear spherical aberration $\Delta f/h_1^2$, where h_1 is the radius of the zone where the rays enter the first lens, as a function of h_1^2, as a function of index of refraction n, and as a function of the spacing of the fourth lens from the third. This spacing permits final adjustment to secure the least possible aberration. Minimum aberration lenses are shown to be more effective in securing zero aberration than aplanatic lenses suggested by Dunoyer.

AT a meeting of the Optical Society of America in 1926 the author described[1] an objective, consisting of four ordinary spectacle lenses ($n=1.523$) each 4 cm in diameter and 2 mm thick, which had practically zero spherical aberration for monochromatic light although the aperture was $f:2.5$. Further calculations show that with a slightly different fourth lens the aperture can be increased to $f:2$ or even more and still keep the diameter of the spherical aberration disk to only a fraction of the diameter of the diffraction disk. Such an objective might be useful in a spectrograph or other camera where chromatic aberration is not important, especially if the light efficiency is doubled by coating the eight surfaces so as greatly to reduce loss by reflection, as is now possible by methods developed since 1926. It is thought, therefore, that a more precise description of this type of objective, and of how the spherical aberration can be made negligible, may be of interest.

Spherical aberration of a thin lens. We shall write the equation for the converging power of a lens as follows:

$$\varphi=\pi'-\pi=(n-1)(\rho-\rho'), \qquad (1)$$

where φ, π, π', ρ, and ρ' denote the reciprocals of focal length f, object distance u, image distance u', and lens surface radii r and r', respectively. All curvatures concave to the right and all distances measured from the lens to the right are taken to be positive.

The reciprocal spherical aberration of a thin lens for rays going through a zone at a radial distance h from the axis, that is the increase in the convergence of these rays in comparison with that of axial rays, for which h is zero, is, in our notation[2]

$$\Delta\left(\frac{1}{u'}\right)=\Delta\pi'=\tfrac{1}{2}h^2\frac{(n-1)}{n^2}\{(\pi-\rho)^2[(n+1)\pi-\rho]$$
$$-(\pi'-\rho')^2[(n+1)\pi'-\rho']\}. \qquad (2)$$

Coddington discovered[3] over a hundred years ago that by introducing a shape function s, which is a function of ρ and ρ' alone, and a position function p, which is a function of π and π' alone, the expression for spherical aberration can be put into a more convenient form. That is, using the following functions:

(shape function)

$$s=\frac{\rho+\rho'}{\rho-\rho'}=\frac{2(n-1)\rho}{\varphi}-1=\frac{2(n-1)\rho'}{\varphi}+1; \qquad (3)$$

[1] G. S. Fulcher, "Dunoyer meniscus lenses for use in spectrographs," abstract in J. Opt. Soc. Am. 12, 476 (1926).

[2] T. Townsend Smith in "Spherical aberration of thin lenses" (Sci. Pap. Bur. Stand. 18, 559–584; Paper No. 461; 1923) gives a derivation of this equation.

[3] H. D. Coddington, *A Treatise on the Reflection and Refraction of Light* (Simkin & Marshall, London, 1829), p. 110. Coddington used a different notation and did not name the functions, but gives equations corresponding to our Eqs. (2) to (8).

(position function)

$$p=\frac{\pi+\pi'}{\pi-\pi'}=-\frac{2\pi}{\varphi}-1=-\frac{2\pi'}{\varphi}+1; \qquad (4)$$

we can transform Eq. (2) into the following equation:

$$\Delta\pi'=\tfrac{1}{8}h^2\varphi^3\left\{\frac{(n+2)}{n(n-1)^2}s^2+\frac{4(n+1)}{n(n-1)}sp+\frac{(3n+2)}{n}p^2+\frac{n^2}{(n-1)^2}\right\}; \qquad (5)$$

and this may be written

$$\Delta\pi'=\tfrac{1}{8}h^2\varphi^3(as^2+bsp+cp^2+d); \qquad (6)$$

where the coefficients a, b, c, and d depend only on n.[4] If we denote (1/8) times the function in parentheses by S,[5] then

$$\Delta\pi'=h^2\varphi^3S;$$

and since $\Delta\pi'=-(\Delta f/f^2)$,

$$S=\tfrac{1}{8}(as^2+bsp+cp^2+d)=-(\Delta f/f)/(h^2/f^2); \qquad (7)$$

and

$$\varphi S=-(\Delta f/h^2). \qquad (7a)$$

Minimum aberration lenses. Partial differentiation of Eq. (6) by s gives as a condition for minimum aberration:

$$2as+bp=0;$$

hence for a minimum aberration lens the shape function is

$$s_m=-(b/2a)p=-\frac{2(n^2-1)}{(n+2)}p. \qquad (8)$$

Substitution of this value in Eqs. (3) gives the corresponding lens surface curvatures:

$$\frac{\rho_m}{\varphi}=\frac{(n+2)-2(n^2-1)p}{2(n+2)(n-1)};$$
$$\frac{\rho_m'}{\varphi}=-\frac{(n+2)+2(n^2-1)p}{2(n+2)(n-1)}; \qquad (9)$$

and substitution in Eq. (7) gives

$$S_m=\frac{n^2}{8(n-1)^2}-\frac{np^2}{8(n+2)}. \qquad (10)$$

This is, of course, the algebraic minimum, taking the sign into consideration, not the absolute minimum.

Aberration due to several thin lenses in contact. We shall distinguish successive lenses by subscripts which indicate their order; then for two lenses in contact, since $\Delta\pi_2=\Delta\pi_1'=h_1^2\varphi_1^3S_1$,

$$\Delta\pi_2'=h_1^2\varphi_1^3S_1+h_2^2\varphi_2^3S_2; \qquad (11)$$

and for four thin lenses in contact

$$\Delta\pi_4'=h_1^2\varphi_1^3S_1+h_2^2\varphi_2^3S_2+h_3^2\varphi_3^3S_3+h_4^2\varphi_4^3S_4. \qquad (12)$$

These equations are not exact since they neglect the effect of the thickness of the lenses, but for thin lenses they are close approximations.

Telescope doublet consisting of two thin minimum aberration lenses with the same index of refraction, in contact.

For the first lens:

$$\pi_1=0; \quad \pi_1'=\varphi_1; \quad p_1=-1.$$

For the second lens:

$$\pi_2=\varphi_1; \quad \pi_2'=\Phi_2=\varphi_1+\varphi_2; \quad p_2=-(2\varphi_1+\varphi_2)/\varphi_2.$$

Assuming $h_1=h_2=h$, we obtain, by use of Eqs. (10) and (12),

$$\Delta\pi_2'=\frac{nh^2}{8(n-1)^2(n+2)}[n(n+2)(\varphi_1^3+\varphi_2^3)-(n-1)^2(\varphi_1^3p_1^2+\varphi_2^3p_2^2)].$$

Substituting Φ_2 for $(\varphi_1+\varphi_2)$ and inserting the values of p_1 and p_2 we obtain

$$\Delta\pi_2'=\frac{nh^2}{8(n-1)^2(n+2)}\times[(4n-1)\Phi_2^3-(2n+1)^2\Phi_2\varphi_1\varphi_2]; \qquad (13)$$
$$=(K_1-K_2\varphi_1\varphi_2)\Phi_2;$$

where K_1 and K_2 are positive quantities whose values depend only upon Φ_2^2, h^2, and n. Evidently the value of $\Delta\pi_2'$ will be a minimum when that of $(\varphi_1\varphi_2)$ is a maximum; and since the product of

[4] I. C. Gardner in "Application of the aberration equation to optical design" (Sci. Pap. Bur. Stand. No. 550, 1927) gives tables of values of the coefficients for values of n from 1.40 to 1.75.

[5] The tables of values of S given by Smith (footnote 2) contain a number of errors.

two quantities whose sum is constant is a maximum when the two are equal, the aberration of the doublet will be a minimum when $\varphi_1=\varphi_2=\frac{1}{2}\Phi_2$. From this it follows that $p_2=-3$ and that

$$\Delta\pi_2'=h^2\Phi_2{}^3\frac{n(2n-1)(5-2n)}{32(n-1)^2(n+2)}. \quad (14)$$

Telescope triplet consisting of three thin minimum aberration lenses with the same index of refraction, in contact. In this case also, assuming h is the same for all lenses, we can prove that for minimum aberration all lenses will have the same focal length. Putting φ_1, φ_2, and φ_3 each equal to $\frac{1}{3}\Phi_3$, we obtain

$$p_1=-1;\quad p_2=-3;\quad p_3=-5;$$

and

$$\Delta\pi_3'=nh^2\Phi_3{}^3\frac{n(8n-5)(7-4n)}{216(n-1)^2(n+2)}. \quad (15)$$

The aberration is evidently zero for n equal to 1.75; hence, by using glass of this index a telescope triplet can be made which will have practically no spherical aberration.

Telescope quadruplet consisting of four thin minimum aberration lenses with the same index of refraction, in contact. In the same way it can be proved that if h is practically the same for all the lenses, the quadruplet will have minimum aberration if the focal length of each lens is equal to $\frac{1}{4}\Phi_4$; then

$$p_1=-1;\quad p_2=-3;\quad p_3=-5;\quad p_4=-7;$$

and

$$\Delta\pi_4'=h^2\Phi_4{}^3\frac{n(10n-7)(2n-3)}{128(n-1)^2(n+2)}. \quad (16)$$

This is equal to zero if $n=1.50$. It is possible, then, by combining four thin lenses made of ordinary crown glass such as is used for spectacles,

TABLE I. Data for quadruplet consisting of four thin minimum aberration lenses of 2.5 diopters each, with $n=1.523$.

Lens	p	s_m	$(n-1)\rho_m$ diop.	$(n-1)\rho_m'$ diop.	S_m	$h^2/h_1{}^2$
1st	−1	0.7491	2.186	−0.314	+1.0060	1.000
2nd	−3	2.2473	4.059	+1.559	+0.5735	.985
3rd	−5	3.7455	5.932	+3.432	−0.2912	.955
4th	−7	5.2436	7.805	+5.305	−1.5883	.915
				Algebraic sum:	−0.3000	

TABLE II. Data for crown glass quadruplet selected for accurate determination of spherical aberration.

Lens	p	$(n-1)\rho$ diop.	$(n-1)\rho'$ diop.	s	S	$(h/h_1)^2S$
1st	−1	2.5	0	1.0	+1.0725	+1.0725
2nd	−3	4.0	1.5	2.2	+0.5760	+0.5674
3rd	−5	6.0	3.5	3.8	−0.2878	−0.2750
4th	−7	7.5	5.0	5.0	−1.5263	−1.3965
				Algebraic sum:	−0.1656	−0.0316

to obtain a telescope objective with little or no spherical aberration.

Crown glass telescope quadruplet with 10 cm focal length and practically zero spherical aberration. We shall assume that $n=1.523$, the value for a commercial crown glass used in eyeglasses; then the equation for S becomes

$$(n=1.523)\quad S=1.0571\cdot s^2+1.5838\cdot ps+0.5392\cdot p^2+1.0600. \quad (17)$$

If the converging power of each lens is taken to be 2.5 diopters then the values of p, s_m, and S_m come out as given in Table I. The sum of the values of S_m for the four lenses comes out −0.30 instead of zero because n is 1.523 instead of 1.50. To correct for the variation of h we shall use the following modification of Eq. (12) for four thin lenses of the same focal length:

$$\Delta\pi_4'=(h_1{}^2\Phi_4{}^3/64)\{S_1+(h_2/h_1)^2S_2+(h_3/h_1)^2S_3+(h_4/h_1)^2S_4\}. \quad (18)$$

The expression in brackets is a sum of weighted values of S. Using the values of $(h^2/h_1{}^2)$ given in the last column of Table I,[6] the algebraic sum of the weighted values comes out −0.160; and since $\Delta f=-\Delta\pi_4'/\Phi_4{}^2$, this means that $\Delta f/h_1{}^2$ comes out only 0.160/640 or 25×10^{-5}/cm. Moreover, the effect of the thickness of the lenses, which we have neglected, would be to bring this value even closer to zero; and, if desired, the value can be further reduced by slight modification of the curvatures of one of the lenses.

Since commercial eyeglass lenses are furnished with curvatures corresponding to integral diopters, halves, and quarters, the lenses listed in Table II were selected as sufficiently close approximations to the lenses indicated by the results in Table I. For these lenses the algebraic sum of

[6] The values of h/h_1 can be determined approximately either graphically or by calculation of the path of a ray through the lenses.

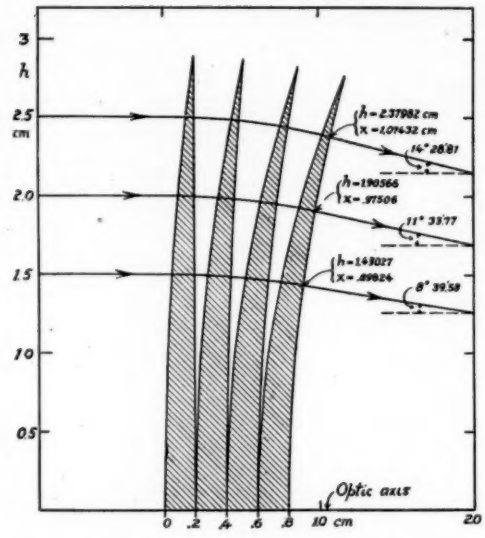

FIG. 1. Section through halves of the lenses of the crown glass quadruplet of 10-cm focal length with zero aberration for apertures up to $f:2$.

the values of $(h/h_1)^2S$ is only -0.032 so that $\Delta f/h_1^2$ comes out only 5×10^{-5}/cm. This is so near to zero that it was decided to determine the aberration accurately for apertures up to $f:2$ for this quadruplet, assuming that the thickness of each lens is 0.2 cm and that the lenses are in contact.

Spherical aberration of the crown glass quadruplet as a function of aperture. Sections through halves of the lenses are shown in Fig. 1. The radii of the lens surfaces, as determined from the figures in columns three and four of Table II, are as follows, in centimeters: 1st, 20.92 and ∞; 2nd, 13.075 and 34.8667; 3rd, 8.7167 and 14.9429; 4th, 6.9733 and 10.46.

We may write Eq. (18) as follows:

$$\Delta\Phi_4=h_1^2\Phi_4^3S_Q=-\Delta f/f^3, \tag{19}$$

where f is the paraxial focal length of the quadruplet. Then

$$S_Q=-(\Delta f/f)/(h_1^2/f^2)=-f(\Delta f/h_1^2);$$
$$\Phi_4S_Q=-\Delta f/h_1^2. \tag{20}$$

Equation (20) gives the relation between the reciprocal spherical aberration coefficient S_Q and the linear spherical aberration coefficient $(\Delta f/h_1^2)$.

(a) Values of $\Delta f/h_1^2$ for large apertures. The paths of rays for h_1 equal to 1.5, 2.0, and 2.5 cm were calculated, using six-place tables; that is, for each ray the coordinates of the points where it met the eight lens surfaces were computed and also the eight angles which it made with the axis after refraction at these surfaces and the positions of the eight image points where the ray would have intersected the axis if extended. The values of $(-\Delta f/h_1^2)$ so obtained are shown plotted as a function of h_1^2 in Fig. 2.

(b) Values of $\Delta f/h_1^2$ for paraxial rays. Figure 2 shows a point for paraxial rays, that is for h equal to zero. The value of $(-\Delta f/h^2)$ for these rays was computed as follows:

Using π_i and π_i' to denote the convergences of the light inside the lens just after refraction by the first surface and just before refraction by the second surface, respectively, we have

$$n\pi_i=\pi+(n-1)\rho; \text{ and } \pi'=n\pi_i'-(n-1)\rho'. \tag{21}$$

If the corresponding distances to the focal points are denoted by f_i and f_i', and the thickness of the lens is denoted by t, then

$$f_i=f_i'+t; \quad \Delta f_1=\Delta f_i'; \quad \Delta\pi_i/\Delta\pi_i'=(f_i'/f_i)^2;$$

and since for rays near the axis, $(h/h')=(f_i/f_i')$, evidently

$$\left(\frac{\Delta\pi_i}{h^2}\right)\Big/\left(\frac{\Delta\pi_i'}{h'^2}\right)=(f_i'/f_i)^4=(\pi_i/\pi_i')^4.$$

Correcting for the change in h due to the distance between the two surfaces, the equation for the reciprocal aberration of a lens becomes, instead of Eq. (2),

$$\Delta\pi'=\tfrac{1}{2}h^2[(n-1)/n^2]\times\{(\pi_i'/\pi_i)^4(\pi-\rho)^2[(n+1)\pi-\rho]-(\pi'-\rho')^2[(n+1)\pi'-\rho']\}. \tag{22}$$

This may be abbreviated to the following:

$$\Delta\pi'/h^2=\tau^4F(\rho,\pi)-F(\rho',\pi'), \tag{23}$$

where

$$F(\rho,\pi)=[(n-1)/2n^2]\times(\pi-\rho)^2[(n+1)\pi-\rho]=\tfrac{1}{2}(\pi-\rho)^2(\pi-\pi_i/n); \tag{24}$$

$$F(\rho',\pi')=[(n-1)/2n^2]\times(\pi'-\rho')^2[(n+1)\pi'-\rho']=\tfrac{1}{2}(\pi'-\rho')^2(\pi'-\pi_i'/n);$$

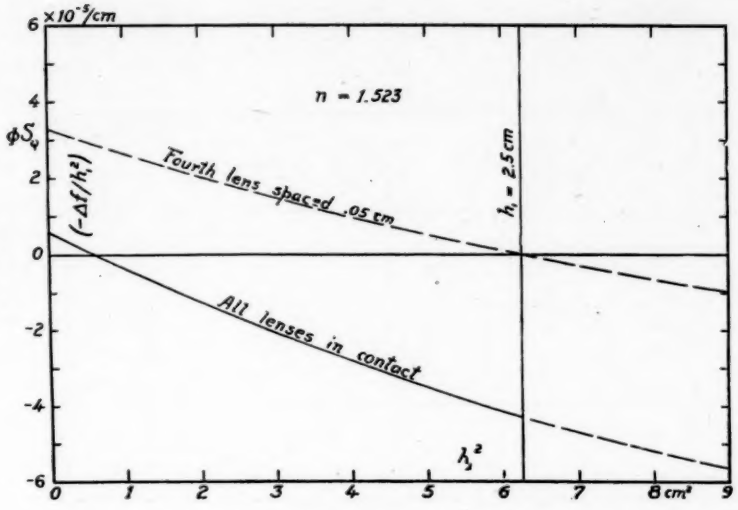

FIG. 2. Plot of ΦS_Q as a function of h_1^2 for n equal to 1.523, (1) with all lenses in contact and (2) with the fourth lens spaced by 0.05 cm.

and

$$\tau = (\pi_i'/\pi_i) = 1 + t/f_i'. \tag{25}$$

For four lenses in contact

$$\begin{aligned} -\Delta f/h^2 = f^2\{&(\tau_1\tau_2\tau_3\tau_4)^4 F(\rho_1, \pi_1) \\ &-(\tau_2\tau_3\tau_4)^4[F(\rho_1', \pi_1') - F(\rho_2, \pi_2)] - \cdots \\ &-\tau_4^4[F(\rho_3', \pi_3') - F(\rho_4, \pi_4)] - F(\rho_4', \pi_4')\}. \end{aligned} \tag{26}$$

Calculation of $\Delta f/h^2$ by use of Eq. (26) involves no trigonometric functions, hence values accurate to eight places of decimals were obtained by use of a calculating machine. For n equal to 1.523 the value of $\Delta f/h^2$ happens to come out only -0.6×10^{-5}/cm, much nearer to zero than the value 5×10^{-5}/cm computed from the data in Table II. This is caused by the effect of the thickness of the lenses in increasing the algebraic value of S_Q.

The value of $\Delta f/h_1^2$ for the large apertures is, however, more important, of course; and as Fig. 2 shows, the value for h_1 equal to 2.5 cm is about 4.3×10^{-5}/cm which gives for Δf the value 27×10^{-5} cm. The corresponding value of the minimum diameter of the image disk due to spherical aberration for aperture $f:2$, however, is only 5×10^{-5} cm. This is practically zero compared with the diameter of the diffraction disk, which for $f:2$ is 4.9 wave-lengths. Nevertheless the diameter of the disk can be reduced to less than 1×10^{-5} cm by spacing the fourth lens 0.05 cm from the third and thus shifting the curve giving ΦS_Q as a function of h_1^2 up to the position shown by the dashed curve in Fig. 2.

Spherical aberration of the crown glass quadruplet as a function of index of refraction. Figure 3 shows how ΦS_Q varies with n both for paraxial rays and for h_1 equal to 2.5 cm. For the paraxial rays the values were calculated by use of Eq. (26); the values for h_1 equal to 2.5 cm were determined from the value for n equal to 1.523 by a differentiation process, that is by determining the changes in angles and in coordinates caused by a given change in index of refraction.

If we assume that the range of index which corresponds to the range of visible wave-lengths is from 1.520 to 1.530, then the spherical aberration is evidently very small for all visible light. In fact, the maximum aberration disk has a diameter of only 7.3×10^{-5} cm. Even this small aberration can be reduced over 40 percent by spacing the fourth lens 0.05 cm from the third, thus shifting the curve for ΦS_Q as a function of n for h_1 equal to 2.5 cm up to the position indicated in Fig. 3 by the dashed curve.

Effect on the spherical aberration of spacing the fourth lens away from the third. The primary effect of moving the fourth lens away from the third is to increase π_4 and $(-p_4)$ and to decrease h_4. The net effect is to decrease the value of $(-h_4^2 S_4)$ and hence to increase the algebraic value of S_Q. Calculations for paraxial rays and for rays with h_1 equal to 2.5 cm gave the results for ΦS_Q plotted

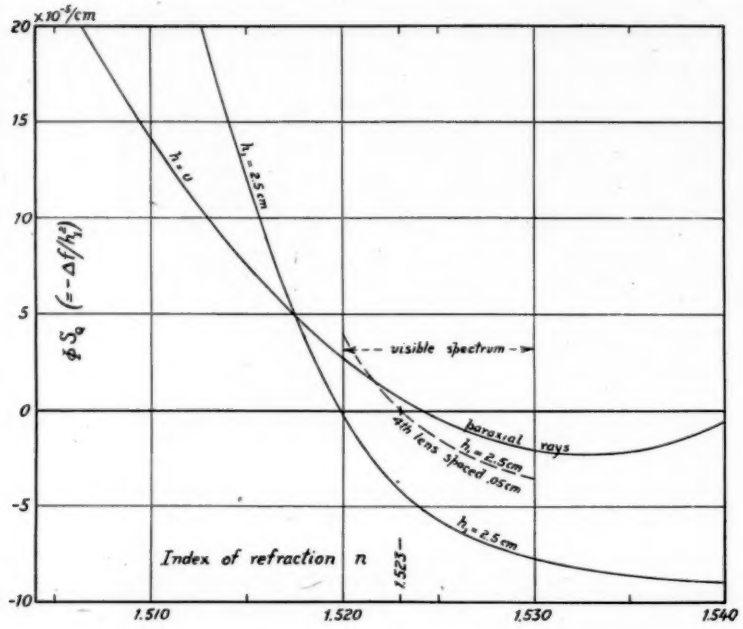

FIG. 3. Plot of ΦS_Q as a function of index of refraction, (1) for paraxial rays and (2) for rays for which $h_1 = 2.5$ cm.

in Fig. 4 as a function of spacing of the fourth lens. As stated above, these results show that by spacing the fourth lens 0.05 cm from the third the value of $\Delta f/h_1^2$ for n equal to 1.523 and for h_1 equal to 2.5 cm can be reduced to zero. For values of n ranging from 1.520 to 1.530 the effect of such a spacing would be to shift the curve for ΦS_Q for h_1 equal to 2.5 cm upward to the dashed curve shown in Fig. 3, and to reduce the maximum diameter of the aberration disk for this range of n to about 4×10^{-5} cm. This crown glass quadruplet, then, has practically zero spherical aberration for apertures up to and even greater than $f:2$ and for a range of index of refraction corresponding to the range of wave-lengths of the visible spectrum.

Use of an objective of large aperture ($f:2$) and of zero spherical aberration in a spectrograph. Let us suppose that the spectrograph has an achromatic telescope lens of focal length f_T and a slit of width w_s and that the crown glass quadruplet of focal length f_Q and of aperture $f:2$ is used as the camera lens. The width of a spectrum line due to diffraction and slit width would be

$$w_L = (k + w_s/f_T)f_Q, \tag{27}$$

where kf_Q, the width due to diffraction, is equal to $(1.22\lambda/h_1)f_Q$. If we make the slit width equal to mkf_T, then

$$w_L = (1+m)kf_Q. \tag{28}$$

It can readily be shown that in this case the intensity of the image per unit area would be independent of the focal length of the telescope lens and proportional to $[m/(m+1)f_Q^2]$. With the quadruplet of aperture $f:2$ used as camera objective, therefore, the intensity of the spectrum lines per unit area would be 16 times the intensity with a camera objective of aperture $f:8$ for the same resolving power. This assumes no extra loss due to reflection, of course.

Since the camera objective is not corrected for chromatic aberration, the photographic plate would have to be accurately placed at the proper angle to the optic axis and even bent slightly around the proper curve to obtain best definition. Figure 5 gives the focal length for paraxial rays for values of n from 1.520 to 1.530. This is measured from the second surface of the last lens.

Dunoyer objective of large aperture. The author's attention was called to the type of objective discussed above in 1915 while doing some spectrographic research at the University of

Wisconsin, by an article by Dunoyer.[7] In this article under the heading "Objective of large aperture without spherical aberration" he described an objective consisting of a minimum aberration lens and two "aplanatic" meniscus lenses, all of index 1.652 but not in contact. An aplanatic meniscus lens is one with ρ/ρ' equal to $(n+1)/n$ and consequently with shape function s equal to $(2n+1)$. In the case of the two meniscus lenses, instead of securing minimum aberration by varying s as we did, Dunoyer kept s constant and varied the position function p by spacing the lenses. If we differentiate the equation for S, Eq. (7), by p we find that for minimum S, $2cp_m=bs$; hence for an aplanatic lens,

$$(s=2n+1) \quad p_m=-\frac{2(n+1)(2n+1)}{(n-1)(3n+2)}. \quad (29)$$

For both the 2nd and 3rd lenses of the triplet, Dunoyer takes s equal to 4.304 and p equal to -5.0334. Assuming the effective focal length of the triplet is 10 cm, this requires that the second lens be spaced 8.847 cm and the third lens be spaced 10.662 cm from the first, making the total distance from first lens to image 21.99 cm. Other data for this objective are given in Table III.

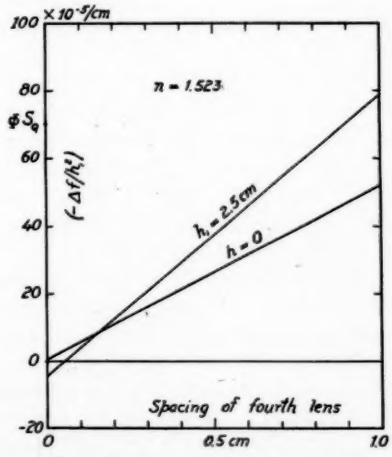

FIG. 4. Plot showing the effect on ΦS_Q of spacing the fourth lens away from the third (1) for paraxial rays and (2) for rays for which $h_1=2.5$ cm.

[7] L. Dunoyer, "On spherical aberration in objectives," J. de physique [5] **3**, 468–486 (1913).

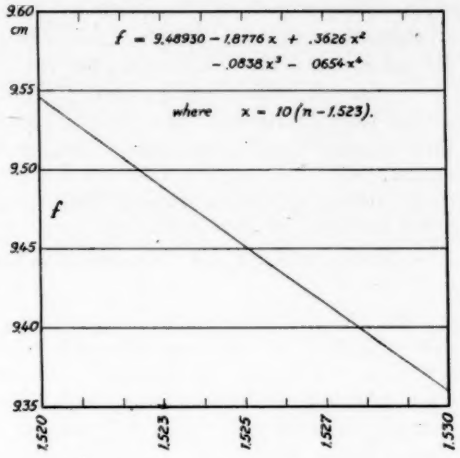

FIG. 5. Focal length of the crown glass quadruplet with all lenses in contact as a function of index of refraction, for paraxial rays.

While the algebraic sum of φ^3S for the three lenses is practically zero, being only $-0.0076\times10^{-5}/\text{cm}^3$, the algebraic sum of $(h/h_1)^2$ times φ^3S for the three lenses is $+1.009\times10^{-5}/\text{cm}^3$. The corresponding result for the crown glass quadruplet is $-0.0316/64{,}000$ or about $-0.05\times10^{-5}/\text{cm}^3$. Evidently the spherical aberration of the Dunoyer triplet is far from zero. Dunoyer made the mistake of assuming that the effect of the change in h is not important. Actually, as we have seen, a triplet with zero aberration cannot be made of glass with an index of less than 1.75; moreover, spacing the lenses increases the algebraic value of the reciprocal aberration of the objective.

Although the article by Dunoyer did not describe an objective without spherical aberration, it stimulated me to work out an answer to the

TABLE III. Data for Dunoyer triplet, with index 1.652 and effective focal length of 10 cm.

	First lens	Second lens	Third lens
s:	0.947	4.304	4.304
p:	−1.0	−5.033	−5.033
S:	+0.746	−0.491	−0.491
100φ:	2.707 diop.	1.765 diop.	2.916 diop.
h/h_1:	1.000	0.7605	0.5082
$10^5\times\varphi^3S$:	+1.4798	−0.2700	−1.2174
$10^5(h/h_1)^2\varphi^3S$:	+1.4798	−0.1562	−0.3144

problem. Dunoyer gave the equation for reciprocal aberration, our Eq. (2) above; but since he did not mention the Coddington functions, I had the fun of rediscovering what Taylor[8] calls "this beautiful device" about 86 years after the publication of Coddington's book.[9] The theory of the crown glass quadruplet was worked out at that time, and two sets of lenses were obtained from the American Optical Company; however, other work prevented a thorough test of the quadruplets. Most of the calculations reported above were made in 1926 while I was with the research laboratory of the Corning Glass Works.

[8] H. D. Taylor, *A System of Applied Optics* (The Macmillan Company, New York, 1906), p. 66.

[9] The functions I called R and P (see footnote 1) are equal, respectively, to $\frac{1}{2}s$ and $-\frac{1}{2}p$. The values of both R and P come out positive for a converging lens forming a real image, and the factor one-half reduces the factor in the equation for S from $\frac{1}{8}$ to $\frac{1}{4}$.

Talbot's Bands

F. B. PIDDUCK
Corpus Christi College, Oxford, England
(Received August 10, 1946)

1.

H. F. TALBOT, the distinguished pioneer in photography, announced in May, 1837 the discovery of the bands that bear his name.[1] These bands are seen when a continuous spectrum is observed through a circular hole the size of the pupil of the eye, with a very thin film of glass over one half. Experiment and theory were alike defective, since Talbot did not state that the bands are only seen when the glass sheet is on the red side, but attributed them to a simple retardation by the glass that would explain them on either side. The complete effect was reported by Brewster[2] four months after Talbot. Airy[3] explained the bands by supposing that each half of the circle produces the Fresnel diffraction pattern of a straight edge, the two patterns interfering with Talbot's retardation. This explains why the bands are only seen on one side. Airy's second paper replaced the semicircles of glass and air by two slits, and investigated their Fraunhofer diffraction. The semicircles were subsequently discussed by Struve.[4]

These theories are the best that could be done in the state of diffraction theory at the time. The diffraction of a glass sheet is caused by the Hertzian oscillators contained in it,[5] which alone can explain plausibly how light can change its direction, the secondary waves of Huygens and Fresnel being entirely non-existent. The diffracted light is indeed expressed as an integral, but over that part of the obstacle that is occupied by matter, not over the aperture. This theory is applied here to Talbot's bands; but I should make it clear at the outset that the matter is one of principle: Struve's theory is nearly correct numerically since the loss of light by reflection in the glass, allowed for in the correct theory, is only 4 percent, and the main effect is that of retardation as assumed by Talbot and Airy.

[1] H. F. Talbot, Phil. Mag. **10**, 364 (1837).
[2] Sir D. Brewster, Brit. Ass. Rep., Notices and Abstracts (1837), p. 12.
[3] G. B. Airy, Phil. Trans. **130**, 325 (1840); **131**, 1 (1841).
[4] H. Struve, Mém. Acad. Sc. St. Petersburg **31**, No. 1 (1883).
[5] F. B. Pidduck, Phys. Rev. **68**, 142 (1945); *Currents in Aerials and High Frequency Networks* (Clarendon Press, Oxford, 1946).

2.

The theory of diffraction by a small thin disk of homogeneous dielectric (conducting or otherwise) may be summarized as follows, as far as it is necessary to consider it here. Take an origin O in the top face of the disk, which occupies the plane of x, y. Let light from a source P_0 fall nearly normally on the disk at a point $P'(x', y', 0)$ and be diffracted at a small angle ψ to a point P as in Fig. 1. Let the electric intensity in the incident wave at distance r_0 from P_0 (the polarization need not be considered here) be $A\exp(i\omega t-ikr_0)/r_0$, where $c=2.998\times10^8$ m/sec. is the velocity of light in a vacuum and $k=\omega/c=2\pi/\lambda$, where λ is the wave-length in meters. Then the electric intensity at P of the Hertzian oscillators in the disk is

$$E=-KA(1-f), \tag{1}$$

where

$$K=\frac{i}{\lambda\times r_0}\int\exp(i\omega t-ikr-ikr_0)dS \tag{2}$$

is a surface integral over the disk with $dS=dx'dy'$ an element of area at P', and

$$f=\frac{\exp(ikb)}{(n+1)^2\exp(inkb)-(n-1)^2\exp(-inkb)} \tag{3}$$

is the fraction of electric intensity transmitted by an infinite sheet of the same material and thickness. In these formulae we have used complex quantities, of which the real parts can be taken in a physical interpretation. We call K Kirchhoff's integral, since it was used by Kirchhoff in his formulation of Huygens's principle. The present theory holds only when the wave-length is much smaller than the disk. If we have, instead of a small disk, a small hole in an opaque sheet, the total E in the diffraction is the E due to the disk, with sign reversed.

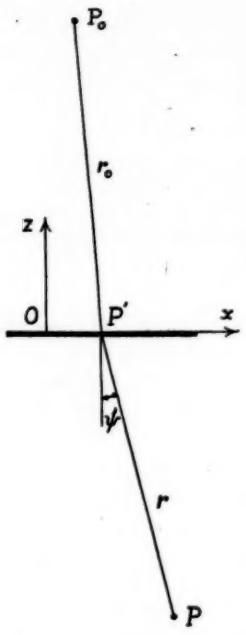

FIG. 1. Diffraction by a plane sheet.

3.

Kirchhoff's integral for Fraunhofer diffraction by a semicircle was calculated by Struve in the special case wanted here. It contains Bessel's function $J_1(z)$ and a function $H_1(z)$ introduced by Struve (see below). Let light fall normally on an infinite opaque sheet S containing a circular hole of radius a, of which the left-hand side is covered with a glass disk of thickness b (Fig. 2). The origin O is at the center of the circle. Let the axis Oz be normal to the sheet and Oy be the straight edge of the glass semicircle, the variable point P' in Kirchhoff's integral having polar coordinates ρ, θ in the plane of x, y. Then for diffraction at an angle ψ in a direction perpendicular to the edge, we have for the glass semicircle

$$K=\frac{i\exp i\omega t}{\lambda r}\int_0^a\int_{-\pi}^{+\pi}\exp(ik\rho\sin\psi\cos\theta)\rho d\rho d\theta.$$

In this formula r_0 has been dropped and A taken to be the amplitude of the electric intensity in the incident wave at the disk. An unimportant phase $-kr_0$ has been dropped. Thus we can take

$$K=\frac{ika^2\exp i\omega t}{2r}\,\frac{J_1(z)-iH_1(z)}{z}$$

for the glass semicircle. For a complete circle

$$K=\frac{ika^2\exp i\omega t}{r}\,\frac{J_1(z)}{z},$$

where

$$z=ka\sin\psi. \tag{4}$$

The electric intensity E of the light diffracted at a small angle ψ by the glass and perforated sheet together is found from Eqs. (1) and (3), remembering Babinet's principle as enunciated at the end of Article 2. We have

$$\frac{E}{A}=\frac{ika^2\exp i\omega t}{2r}\,\frac{(1+f)J_1(z)+i(1-f)H_1(z)}{z}. \tag{5}$$

$H_1(z)$ is an even function of z, tabulated from $z=0$ to 16 in Watson's *Bessel Functions*, pages 667–97. Other properties of the function, wanted below, are found on pages 328–9 of the same book.

If the refractive index n of the glass sheet is 1.5,

$$24/f=25\exp(\tfrac{1}{2}ikb)-\exp[-(5/2)ikb]. \tag{6}$$

The retardation of phase due to the glass being $R=(n-1)kb=\frac{1}{2}kb$ in radians, f is then periodic in R with period 2π. Separating f into its real and imaginary parts $\alpha+i\beta$, the intensity of the light

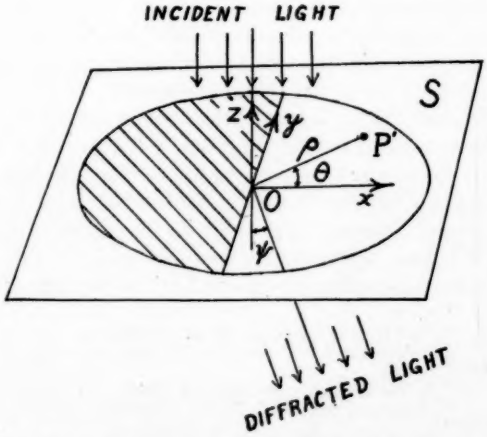

FIG. 2. Diffraction by a glass disk half covering a circular hole.

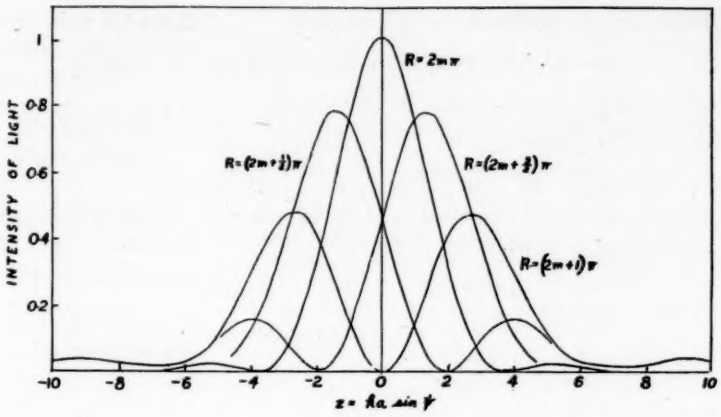

FIG. 3. Fraunhofer diffraction by a glass disk of refractive index 1.5, half covering a circular hole.

diffracted at an angle ψ in Fig. 2 is given by

$$J/J_0=\{AJ_1^2(z)+2HJ_1(z)H_1(z)+BH_1^2(z)\}/z^2, \quad (7)$$

where

$$A=(1+\alpha)^2+\beta^2, \quad H=2\beta, \quad B=(1-\alpha)^2+\beta^2, \quad (8)$$

and J_0 is the intensity at diffraction angle 0 when $R=2m\pi$. Figure 3, plotted from these formulae, shows J/J_0 for various angles when $R=2m\pi$, $(2m+\frac{1}{2})\pi$, $(2m+1)\pi$, and $(2m+\frac{3}{2})\pi$. The curve for $R=2m\pi$ is the same as the curve of intensity in Fraunhofer diffraction by a circular aperture in an opaque sheet. The curve for $R=(2m+1)\pi$ is quite different and has a zero of intensity at $\psi=0$. The other curves are unsymmetrical. The first minimum (which is always nearly zero) passes through $\psi=0$ from right to left when R increases through $(2m+1)\pi$. This is the cause of Talbot's bands (Article 4).

4.

These bands can be discussed in various ways. Airy and Struve gave figures for the intensity produced by dispersion and diffraction acting together. It will be simplest here, now that we have explained the physical theory, to confine attention to the main point, the presence of dark bands. These arise from the fact that the shift of the first minimum through $\psi=0$, just referred to, is in the opposite direction to the shift by dispersion if the glass is on the red side, resulting in a band of low intensity which is nearly stationary for a considerable fraction of the interval between $R=2m\pi$ and $2(m+1)\pi$.

To fix ideas, restrict k to the value 1.066×10^7 that it has for sodium light, and let $R=\frac{1}{2}kb$ be exactly $(2m+1)\pi$ for sodium light. Let $k+dk$ be a near wave number corresponding to light of slightly different wave-length, and let $R=(2m+1)\pi+\epsilon$ for it. From Eq. (6), when $n=1.5$,

$$24/f=-25(1+i\epsilon)+(1-5i\epsilon),$$

to the first order in ϵ, and to that order

$$\alpha=-1, \quad \beta=(5/4)\epsilon, \quad A=0, \quad H=(5/2)\epsilon, \quad B=4.$$

If z is small, $J_1(z)/z=\frac{1}{2}$ and $H_1(z)/z=2z/3\pi$ to the order needed here. Equation (7) gives $J/J_0=5\epsilon z/3\pi+16z^2/9\pi^2$ to order 2 in ϵ and z together. Differentiating with respect to z, we have for the first minimum $z=-15\pi\epsilon/32$. Since $R=\frac{1}{2}(k+dk)b$ for the wave number $k+dk$, $\epsilon=\frac{1}{2}bdk$; and z is approximately $ka\psi$ from (4). Hence the angle of diffraction of the first minimum, for a wave number $k+dk$ near to that of sodium light is

$$\psi=-15\pi bdk/64ka=-0.736bdk/ka. \quad (9)$$

Let the prism in Talbot's bands be a 60° prism of glass for which

$$n=1.5, \quad dn/dk=2.4\times10^{-9}$$

for sodium light, an average silica crown. The minimum deviation is $D=0.649$ radian for sodium light and $D+dD$ for wave number $k+dk$, where $dD=1.51dn=3.63\times10^{-9}dk$. The first minimum is stationary under diffraction and dis-

persion together if $dD+\psi=0$, or from (9) if

$$b/a=3.63\times10^{-9}k/0.736=0.0526. \quad (10)$$

This gives the best thickness for Talbot's bands with sodium light. Since $(2m+1)\pi=\frac{1}{2}kb$,

$$2m+1=b/\lambda. \quad (11)$$

Consecutive bands are separated by a wavelength which is nearly

$$d\lambda=\lambda/m. \quad (12)$$

5.

For a numerical example;

If $a=1.5$ mm, corresponding approximately to the pupil of the eye, $b=0.079$ mm, $m=67$, and $d\lambda=88$A.

If $a=1.25$ cm, corresponding approximately to the tube of a spectrometer, $b=0.66$ mm, $m=560$, and $d\lambda=10$A.

If the disk is made of crown glass with $n=1.5$ for sodium light as before, and the prism of a silica flint for which $n=1.6$, $dn/dk=5.2\times10^{-9}$, then

If $a=1.5$ mm, $b=0.19$ mm, $m=160$, and $d\lambda=37$A.
If $a=1.25$ cm, $b=1.57$ mm, $m=1330$, and $d\lambda=4$A.

Book Review

The Testing and Adjustment of Telescope Objectives

H. Dennis Taylor. Pp. 95+viii, Figs. 47. Sir Howard Grubb, Parsons and Co., Newcastle upon Tyne, England, (Fourth edition) 1945. Price 17/6.

This famous little book, originally written in 1891 and reissued in 1896 and 1921, will warm the heart of anyone who has ever looked through a large telescope. Mr. Dennis Taylor has a keen eye, a passion for precise and careful observation, and a delightful way of communicating his enthusiasm to his readers. He writes in the tradition of Rayleigh, with a personal touch that carries conviction and makes the whole subject live.

To Dennis Taylor, optics has always been an art and never a science. With him the craftsman's touch far outreaches dull computations, and the best mechanical work is but a poor substitute for careful hand-finishing and adjustment. The precise balance of colored fringes and diffraction rings, which makes all the difference between a perfect and an imperfect objective, has provided him with ample material for the writing of this little book.

Most of the chapters are devoted to a detailed description of the star test and knife-edge test of a telescope objective, with careful diagrams and reproductions of the appearances characteristic of chromatic, spherical, and zonal aberration, coma, and astigmatism. The effects of decentering and "squaring-on" of an objective, and the contributions of eye and eyepiece to the observed image of a distant star are described in the greatest detail, and the most minute instructions are given for setting up a large telescope for observation. Some interesting remarks are added as to the general treatment of objectives, including such matters as the growth of fungus on lens surfaces in a damp atmosphere, and the beneficial effects on light transmission which can result from tarnish on the glass.

The book concludes with three interesting papers by Dennis Taylor reprinted from the *Monthly Notices* and *Memoirs of the Royal Astronomical Society*, dealing with the effects of secondary spectrum and describing a new, perfectly achromatic refractor (the Cooke Photo-Visual) in which secondary spectrum has been eliminated.

Astronomers and telescope makers are alike indebted to the Grubb-Parsons Company for their enterprise in republishing this valuable book in these difficult times. The typography and general production are excellent, and all of the original illustrations have been carefully reproduced.

R. Kingslake
Eastman Kodak Company

Letter to the Editor

Visual Thresholds of Steady Point Sources of Light in Fields of Brightness from Dark to Daylight*

SELIG HECHT
Laboratory of Biophysics, Columbia University, New York City
October 10, 1946

IN a recent paper with the above title, Knoll, Tousey, and Hulburt[1] present measurements of the brightness i of a tiny spot of light when it first becomes visible against a background whose brightness b is varied between dark and daylight. They point out that the data show an obvious bend "at the transition from foveal to extra-foveal vision." However, as an analytical description of the data they draw a *single* curve whose equation is

$$i=c(1+b)^{\frac{1}{2}} \tag{1}$$

where the constant c equals 10^{-10} when i is in footcandles. They conclude that the equation agrees with the measurements "within a factor of 3 for b ranging over nine orders of magnitude." The data and curve are in Fig. 1.

This paper invites two comments. (1) It may be useful as rough engineering practice to write *one* equation for the whole visual range. However, the resulting deviations of the equation from the data cannot be overlooked. These systematic and gross deviations arise because the meaning of the bend "at the transition from foveal to extra-foveal vision" has been ignored. A similar bend has been found in measurements of brightness discrimination, visual acuity, dark adaptation, instantaneous threshold, and flicker recognition.[2] It stems from the separateness of day vision and night vision in accord with knowledge accumulated from Purkinje's day to the present. The measurements themselves and all relevant knowledge call for *two* curves to describe the whole visual range.

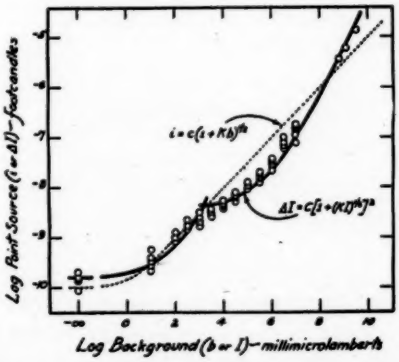

FIG. 1. Measurements of the brightness of a point source when it just becomes visible against a luminous background.

However, even two curves drawn from Eq. (1) cannot describe the data. The equation requires that at very high brightnesses the relation between $\log i$ and $\log b$ have a slope of $\frac{1}{2}$, whereas the data give the slope as almost 1.

In passing, it should be noted that the variable b is without a constant only by chance. In units other than millimicrolamberts it requires a constant K. Therefore the equation is really

$$i=c(1+Kb)^{\frac{1}{2}} \tag{2}$$

where $K=1$ by chance.

(2) The measurements are of brightness discrimination. The small size of the area of augmented brightness does not change the fact that the experiments define the relation between the background brightness b and the just discriminable brightness increment i, a problem first recognized for vision by Bouguer in 1760 and posed for all senses by Weber in 1834.

In 1934, I derived an expression for brightness discrimination which describes all the known measurements.[3] This equation is

$$\Delta I=C[1+(KI)^{\frac{1}{2}}]^2 \tag{3}$$

where, as commonly used, I is the background brightness, and ΔI is the added brightness required to make a given area just perceptible against the background. On a log plot the constants C and K determine the vertical and horizontal position of the curve, and at high brightnesses the slope of the curve is 1.

To cover the data for the whole visual range, *two* pairs of constants have to be used. For the low brightness segment $C=10^{-9.80}$, and $K=10^{-1.90}$; for the high brightness segment $C=10^{-8.35}$, and $K=10^{-5.90}$. The corresponding curves in Fig. 1 fit the measurements not within a factor of 3, but more nearly within the spread of the measurements. Since the measurements were made with the natural pupil, a better agreement between data and equation could be secured by correcting for pupil area including the Stiles-Crawford effect. However, for the present purposes this is not necessary.

The data of Knoll, Tousey, and Hulburt are thus of interest because they show that even for a point source, brightness discrimination follows its usual pattern in visual function. The measurements agree with previous data and theory.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.
[1] H. A. Knoll, R. Tousey, and E. O. Hulbert, J. Opt. Soc. Amer. **36**, 480 (1946).
[2] S. Hecht, "Rods, Cones, and the Chemical Basis of Vision," Physiol. Rev. **17**, 239 (1937).
[3] S. Hecht, "A theoretical basis for intensity discrimination in vision," Proc. Nat. Acad. Sci. **20**, 644 (1934).

Meeting of the Inter-Society Color Council

THE Sixteenth Annual Meeting of the Inter-Society Color Council will be held in New York City on February 24 and 25, 1947, following the meeting of the Optical Society of America. The program, which will be held at the Hotel Pennsylvania, is as follows:

February 24 Hotel Pennsylvania

9:00 A.M. Registration

9:30 A.M. Discussion Session

Based on reports of problems under investigation by Council committees

Problem 6: Color Terms
10: Color Aptitude Test
11: Color Blindness Studies
12: Illuminating and Viewing Conditions for Colorimetry
13: The Illuminant in Textile Color Matching

2:00 P.M. Discussion Session—Continued

3:00 P.M. Business Session

February 25 Commodore Hotel

9:30 A.M. Technical Session on Color held jointly with the Technical Association of the Pulp and Paper Industry

The Paper Man's Interest in Color. A. H. CROUP

Colorimetric Standardization in Industry. RALPH M. EVANS

Color-Order Systems. CARL E. FOSS

From a simple series of standards based on physical types to systems that represent increasing definition of a useful color space.

2:00 P.M. **Spectrophotometry.** NORMAN BARNES

I.C.I. Standard Observer and Coordinate System. J. A. VAN DEN AKKER

Inter-relation of Color Specifications. DOROTHY NICKERSON

Color Engineering. FREDERIC H. RAHR

All who are interested are invited to attend any or all of these sessions.